DOCUMENT RESUME

ED 265 695 EC 181 385

AUTHOR Haring, Norris
TITLE Investigating the Problem of Skill Generalization. 3rd Edition. With Literature Reviews.
INSTITUTION Washington Univ., Seattle. Washington Research Organization.
SPONS AGENCY Department of Education, Washington, DC.
PUB DATE Jul 85
CONTRACT 300-82-0364
NOTE 129p.; For second edition, see ED 249 737. For first literature review, see ED 249 738.
PUB TYPE Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Environmental Influences; *Generalization; Literature Reviews; *Self Control; *Severe Disabilities

ABSTRACT

The report describes second year efforts of a 5-year investigation of skill generalization among severely handicapped students. An overview section introduces the Washington Research Organization, noting its administration and management and its major emphasis on the collection of descritptive data and laboratory research, research in natural educational settings, evaluation, and communication. Basic research concepts in generalization are reviewed. The four approaches to generalization are summarized: (1) studies in ecological variables, (2) studies in performance patterns, (3) studies in self control, and (4) studies in strategy implementation. The second part addresses each of the four generalization strategies by means of research reviews: training in the natural environment and skill generalization, response competition and generalization, teaching retarded students to reinforce their own behavior, and social skills training. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

6072

ED265695

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Norris G. Haring

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Washington Research Organization

Research in Education of the Severely Handicapped
University of Washington

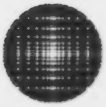
INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION

3rd Edition

with Literature Reviews

Norris Haring
Principal Investigator

EC181385



Washington
Research
Organization

INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION

3rd Edition

Norris Haring
Kathleen Liberty
Felix Billingsley
Owen White
Valerie Lynch
Joan Kayser
and Frances McCarty

with Literature Reviews by

Felix Billingsley
Sharon Field
Carol Kiolet
B. Douglas Leber
Kathleen Liberty
Barbara Matlock
Larry Michael
Claire Phifer
Marsha Thompson
Greg Weisenstein
and Owen White

Research in Education of
the Severely Handicapped

University of Washington
110 Miller Hall, DQ–12
Seattle, WA 98195

EC181 385

3

Norris Haring, Principal Investigator
Kathleen Liberty, Project Coordinator

Michael Boer, Editor

The activity which is the subject of this report was supported in whole or in part by the U. S. Department of Education (Contract No. 300-82-0364). However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Printed in the United States of America.

July 1985

Table of Contents

Table of Contents	**v**
Executive Summary	**vii**
Part 1: Overview	**1**
ONE: The Washington Research Organization	**4**
Research in Generalization	5
Evaluation	6
Communication	7
Advisory Committee	9
Administration and Management	10
TWO: Basic Research Concepts	**12**
Skill Generalization	12
Instructional Programming for Generalization	14
Methodology	16
THREE: UWRO's Approaches to Generalization	**21**
Studies in Ecological Variables	23
Studies in Performance Patterns	38
Studies in Self-Control	45
Studies in Strategy Implementation	54
FOUR: Putting It All Together	**58**
Part 2: Literature Reviews	**61**
FIVE: Training in the Natural Environment and Skill Generalization: It Doesn't Always Come Naturally	**63**
SIX: Response Competition and Generalization	**80**
SEVEN: Teaching Retarded Students to Reinforce Their Own Behavior: A Review of Process and Operation in the Current Literature	**88**
EIGHT: Social Skill Training Review	**107**

Executive Summary

The Washington Research Organization (UWRO), an Institute for Research in Education of the Severely Handicapped, is conducting a five-year investigation of the problem of skill generalization. UWRO is investigating three approaches to developing strategies for facilitating skill generalization; a fourth area of studies is aimed at developing a comprehensive plan for verifying the practical application of the results of the Institute's research in natural educational and vocational settings.

Performance Pattern Studies began with a retrospective analysis of existing data sets, and will proceed to the collection of descriptive data in public school classrooms. These data will be used to determine a set of experimental decision rules for matching specific instructional methods to individual learners.

Ecological Studies initiated a four-year longitudinal descriptive study of factors in educational settings which may influence generalization. Intervention studies are also included within this approach, with studies of massed vs. distributed instructional trial sequencing and the effects of competing behaviors.

Self-Control Studies commenced with studies of the effects of self-monitoring procedures on skill generalization. This area of studies also includes investigations of self-reinforcement and self-instruction.

Strategy Implementation Studies will be carried out with a variety of natural training and community settings. Through these experiments, methods will be developed to combine the results from the Ecology, Performance Pattern, and Self-Control studies into an integrated set of best practices for generalization.

The Performance Patterns, Ecological, and Self-Control areas conducted studies during the first two years, and will continue through the next three years. During the second year, UWRO also supported studies on the effects of cue-fading in the secondary/post-secondary transition. Strategy Implementation studies will begin in the third year.

Performance Patterns

The objective of the Performance Patterns research is to identify critical elements in responding that will improve the accuracy of

predictions of generalization, and combine or match the particular types of instructions with the students' need in order to facilitate generalization.

By examining the performance characteristics of students who are acquiring a new skill or building fluency in a skill, we have learned that certain elements of performance are critical to effective planning for instruction. These include the student's correct rate of performance, the accuracy, the weekly rate of progress, and the variability of performance. By examining those characteristics we have been able to predict if certain strategies will improve the student's performance. We have seen that there is a strong consistent relationship between a pupil's fluency on a task and his/her need for additional instruction.

We have continued to gain valuable information by analyzing the relationship between performance patterns and generalization as reported in published research on generalization. During the first two years, a retrospective analysis of published generalization research was conducted by Owen White. The behaviors studied included the following classifications: social, communication, vocational/prevocational, self-help/independent living, cognitive/academic skills, and cognitive strategies. A total of 115 studies have been evaluated to date. The great majority (93%) of the behaviors studied had immediate functional utility to the subject.

Practically all of the training strategies identified by Stokes and Baer designed to lead to generalization were included in the studies reviewed. The most common strategy was "train and hope." A greater proportion of the more recent research used "loose training," "sufficient exemplars," and "natural maintaining contingencies."

This information has been valuable in providing direction for the Performance Pattern research. The next step will be to study the relationship between those patterns of learning and the influence that they may have in facilitating the generalization of new skills.

In addition to the retrospective analysis of reported studies, experimental studies have continued concentrating on the role of fluency in determining the probability of skill generalization. A second series of studies has been concerned with the impact of natural consequences on skill acquisition and generalization.

It is the purpose of the performance pattern research to study these elements to determine their relationship to generalization. Once a relationship is established, the study will attempt to match specific patterns with instances of generalization—with the anticipation that these patterns will provide a basis for predicting which strategies will facilitate generalization.

Ecological Studies

The Ecological research consists of an ongoing longitudinal descriptive study of ecological variables and a series of studies on individual aspects of the instructional ecology.

The longitudinal descriptive study will continue for a total of four years, pursuing four issues: (a) the extent to which teachers of the severely handicapped actually establish goals and objectives that promote generalization; (b) the extent to which student performance indicates that they have attained goals and objectives involving generalization; (c) the degree to which generalization actually occurs that is a result of programming; and (d) the general ecological conditions which might facilitate generalization.

In order to determine the extent to which teachers are programming for generalization, Felix Billingsley did an analysis of IEP objectives to determine whether or not skill performance was likely to have adaptive value. He found that a very small percentage of the objectives specified generalization intent. Billingsley concluded that teachers are spending considerable time teaching behaviors which have low probability of generalization.

In a second study, parents were asked to rate their child's objectives and opportunities to generalize skills taught in school to the home. Billingsley found that the parents considered the majority (91%) of the skills being taught by the schools to be functional and that their children had the opportunity to perform the majority of these learned behaviors in the home and community. Generally these behaviors were found to occur appropriately within non-training settings. Further, it was found that 85% of the behaviors which parents indicated were being performed appropriately were being trained at home.

A third study was conducted to determine the degree of access the students have to other settings and managers. It was found that

students have a considerable variety of managers during the school day. In the sites observed, the conditions seemed to be favorable to the development of generalized responding across persons; however, most of the students spent a large part of their in-school time in only one setting.

Assuming that nonhandicapped peer interaction is important to the generalization of communication and social skills, a fourth study was conducted to determine the amount of time the subjects had opportunities to interact with nonhandicapped peers. It was found that the amount of time available for interactions in or out of school was very low.

During the second year of the ecological studies, the longitudinal study series was continued and expanded, with continued focus on ecological conditions including manager and setting variations, interaction opportunities, across-setting generalization, and the content of a statewide survey of IEPs.

In addition to the longitudinal study series, two intervention studies were conducted. The first, on the effects of competing behaviors on generalization, explored the possibility that the existence of old behaviors in an individual's repertoire may compete with the generalization of newly acquired target behaviors. The second intervention study investigated the effects of trial sequencing on generalization. This study evaluated the effects of traditional and functional models on the acquisition and generalization of a functional task.

Studies in Self-Control

In self-control training, individuals are taught methods to direct their own behavior. Several procedures are useful in self-management. One is self-monitoring, which involves "keeping track" of one's own behavior. Another procedure is self-reinforcement. This includes choosing the reinforcement and providing it after one has completed the task or restrained from giving in to an undesirable habit. A third procedure is self-instruction. This involves deciding the task to be done, then planning the order best to approach and complete the task, and after a certain amount of trial and error the task may be completed.

Self-monitoring, self-reinforcement, and self-instruction can be powerful procedures leading to a relatively high level of independence for handicapped individuals. In addition, if taught effectively, self-control may exceed the effectiveness of external control in facilitating maintenance and generalization.

Obviously a crucial question is: Can severely handicapped individuals be taught to employ the processes involved in self-control? Based on preliminary exploration, Kathleen Liberty is optimistic about the value of self-control in facilitating generalization. During the first year of UWRO investigations, she conducted a series of three basic studies, the results of which have provided support for her optimism.

The first self-control study was designed to examine: (a) the acquisition of self-monitoring by a severely handicapped student through an avoidance training procedure, and (b) the effects of self-monitoring on the target behavior. Liberty found that the training procedures produced rapid acceleration of independent use of a counter without the addition of specific reinforcement for self-monitoring. Further, the data on the target behavior indicated that self-monitoring acted as a positive reinforcer sufficient to maintain performance of the target behavior.

The second study in the series was designed to: (a) examine the maintenance of self-monitoring skills; (b) determine if self-monitoring generalized within stimulus classes and across responses; and (c) examine the effect of self-monitoring on the target behaviors. Liberty found that self-monitoring maintained at high levels of reliability and independence. In addition, self-monitoring did generalize across different stimulus conditions within the same class, and when the stimulus materials were changed reliable self-monitoring was established. However, independent and reliable self-monitoring did not transfer across behaviors. A self-monitoring avoidance prompting procedure was introduced for assembling sack lunches. A few prompts were sufficient to produce generalized counting at 100% reliability.

During the second phase, wearing a wrist counter produced an increase in rate sufficient enough to have practical value to the subject. The procedure produced a median increase in production to 67.5% of normal and reached as high as 114% of normal. Self-monitoring was found to have a direct affect on increasing the rate of performance on the lunch bagging task.

In a third study, designed to extend and replicate the results of the previous study with different behaviors and a different subject, the subject was instructed to use two-word answers to questions in the training setting only. The uninstructed response was two-word initiation. Data were also collected on actuation of a wrist counter, and the independence and reliability of self-monitoring. Instruction in self-monitoring of two-word answers in the training setting was found to produce rapid acceleration of independent self-monitoring. Liberty concluded that by itself, the wrist counter did not substantially affect responding; however, once training in the wrist counter was initiated, the counter was found to mediate differences in consequation between settings. Further, training self-monitoring was a simple and efficient way to mitigate differences in settings when skill transfer is desired.

During the second years studies, Liberty continued the investigation of questions related to self-monitoring and self-reinforcement procedures. At the end of this series of investigations Liberty concluded that the opportunity to self-monitor improved the level of generalization of the target behavior for two severely handicapped youths. The second study is yet inconclusive because during the period of the study the subjects did not learn to self-reinforce.

Studies in Strategy Implementation

A new series of applied studies was added to implement the most positive results of preceding more basic research in mature instructional and vocational settings. We were interested in testing the effects of generalization-promoting strategies on skill generalization by severely handicapped students when these strategies are implemented.

This Publication

This third edition is an accumulation of the previous editions with specific emphasis on the second year's activities. The three research teams responsible for studies in ecological variables, performance patterns, and self-control headed up by Billingsley, White, and Liberty respectively have proceeded to continue and expand in those areas. The first studies in Strategy Implementation are underway as this publication goes to press.

During the second year of research, UWRO investigators conducted reviews of applicable literature within each of these four general approaches:

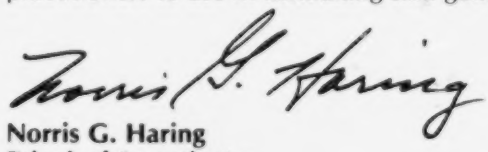
Owen White, Doug Leber, and Claire Phifer present, in "Training in the Natural Environment and Skill Generalization: It Doesn't Always Come Naturally," a descriptive analysis of 115 published studies from UWRO's retrospective analysis of existing data sets.

Barbara Matlock, Felix Billingsley, and Marsha Thompson's "Response Competition and Generalization" provides a review of the literature that pertains to one aspect of UWRO's ecological studies.

Kathleen Liberty and Larry Michael's review of the literature, "Teaching Retarded Students to Reinforce Their Own Behavior: A Review of Process and Operation in the Current Literature" is related to the subject of the UWRO studies in self-control.

Greg Weisenstein, Sharon Field, and Carol Kiolet prepared reviews for their UWRO studies on "Social Skill Training Review." This review may be relevant to determining ways in which peer social interactions may be associated with generalization.

UWRO will produce a total of four literature review products, of which this is the second. In addition, the Performance Patterns, Ecological, Self-Control, and Strategy Implementation research approaches will continue through the next three years, with the ultimate objective of producing a unified set of guidelines for practitioners to use in facilitating skill generalization.



Norris G. Haring
Principal Investigator
Seattle, 1985

Part 1:

Overview

Ms. Cindy Burchart is pleased with her new job at the Seattle Hotel. She has loaded the industrial dishwasher for the first time: all of the plates on the bottom in neat rows and all of the glasses on the top. It was easy to figure out where they went. She closes the door with satisfaction. But where are the buttons to start the machine? They're not on the front of the machine, nor on the side. Behind the dishwasher, on the wall, Cindy sees a row of buttons, switches, dials, and lights. Some of the lights are dark, while others are glowing red or green. She stands bewildered before the display. The manager of the kitchen rushes over, glares at Cindy, and rapidly pushes some buttons, sets a dial, and flicks a switch. He barks, "Start on the next load," wondering why he ever agreed to give a retarded person a chance, anyway.

Mr. White gazes at the assessment data for Richard. He is depressed; this is the third year he has had Richard in his class, the third year he has conducted assessment, and the third year he must prepare instructional objectives for Richard's IEP. Last year he taught Richard to say, "My name is Richard Clark," when asked "What's your name?" or "Who are you?" This year, he only answers with, "Richard." "That really won't help if he gets lost," sighs Mr. White. He ruefully writes the objective for "says own name" for the second year in a row. He looks at some more assessment data, collected over the first six weeks of school. It is taking Richard even longer to learn to say his address than it did to say his name, and it looks like there is no guarantee that he will remember that next year. Mr. White considers just getting him an i.d. bracelet, but remembers what Richard's parents said. He writes an objective for "says own address" and shakes his head.

Jody is screaming so loudly that his face is eggplant purple. Mrs. Loomis stares helplessly at him. She goes over and picks up the tennis shoes from the corner where he threw them moments ago. She knows that Jody's teacher told her that Jody was able to put on these very same shoes without any help. The screaming is now broken by gasps, as Jody winds up to an even higher pitch. Mr. Loomis yells up the stairs, "Where are you? We're all in the car waiting!" Mrs. Loomis quickly picks up Jody, puts his shoes on him, and carries him down stairs. Jody quits screaming when they go out the door. "Thank goodness," she says to herself.

ONE

THE WASHINGTON RESEARCH ORGANIZATION

The problem for Cindy, Richard, and Jody is generalization, or the lack of it, to be more precise. The setting changes, time passes, and it is somehow as if they had never learned what to do in the first place. This is one of the most important problems we have had to face since we began educating severely handicapped students. It is one that must be solved if education is to be truly a preparation for life in society.

The Washington Research Organization (UWRO), located on the campus of the University of Washington in Seattle, was awarded a five year contract by the U.S. Department of Education's Special Education Projects (SEP) in October 1982. The mission of UWRO is to identify instructional strategies through empirical investigations that enable practitioners to promote generalized responding by severely handicapped persons. If the research we conduct is productive, we will develop practical instructional methods that ensure that severely handicapped individuals are able to use learned skills in environments outside of the training setting, and that those skills remain useful long after formal education has ceased. We will also take steps to see that the methods we develop are available to practitioners.

A sequence of objectives will need to be met if we are to be successful in accomplishing our mission. First, UWRO investigators will conduct descriptive and intervention research to identify specific environmental, instructional, and individual characteristics that affect the probability of generalized responding. Although variables so identified may increase our understanding of generalization, little of practical value is accomplished by identification alone. Therefore, our second objective is to conduct research designed to test the effects of manipulating or changing environmental, instructional, and individual performance variables. Third, UWRO investigators will conduct intervention research in controlled and natural settings to determine instructional and curricular strategies that increase the probability of generalized responding. This leads us to our fourth objective. Teachers and other practitioners will be trained to implement the strategies identified by UWRO research, in order to determine whether procedures may be used effectively and practically, within the operating and budgetary constraints of educational settings.

The success of meeting each objective will be determined by changes in pupil performance data, and by determining the overall practical impact of such changes. The effects of interventions will be evaluated according to the change in frequency, quality, and quantity of generalized responding from pre-intervention levels. The results of UWRO's studies will be evaluated according to psychological and educational research standards for reliability, validity, methodological considerations, and analytic techniques.

The extent to which severely handicapped individuals successfully demonstrate generalized behavior will be the extent of our success in meeting these objectives.

The activities of the Washington Research Organization are designed to meet these objectives and are organized around four major activity categories: descriptive and laboratory research, research in natural educational settings, evaluation, and communication. These tasks will be supported by the activities of the Advisory Committee and by project management. This document describes the activities of UWRO, basic concepts in generalization research, and our research approaches to the problems encountered by Cindy, Richard, and Jody.

Research in Generalization

Research in generalization constitutes the major activity of UWRO. These activities are divided into two categories, "Descriptive and Controlled Laboratory Studies" (Task 1) and "Research in Natural Settings" (Task 2). Task 1 activities are designed to identify precise variables that affect generalization and to test specific strategies under tight experimental control. They will be conducted primarily during the first two years. These studies are designed to provide the background information needed so desperately in our understanding of generalization.

Task 2 research will investigate the effects of interventions in natural educational settings. In the fourth project year, Task 2 research activities will include investigations of the efficacy of guidelines developed for practitioners from UWRO research. The guidelines will define how procedures are to be applied in natural settings. These investigations will seek to determine how applying the guidelines affects the generalization demonstrated by severely handicapped individuals, and also how guidelines might be improved for more accurate and effective implementation.

We are fortunate to have established cooperative arrangements with five local school districts to provide the settings and subjects for our research. These local educational agencies are Lake Washington School District No. 414, Northshore School District No. 417, Issaquah School District No. 411, Bellevue School District No. 405, and Tacoma School District No. 10. Personnel from these districts will work closely with the Senior Investigators, providing the first contact with parents, guardians, and teachers and arranging for research settings.

Representatives of the districts comprise the Direct Service Consortium. Ralph Bohannon, Director of Special Services for Lake Washington, is an experienced researcher and has cooperated in previous University of Washington research projects. This large district is also represented by Nancy Wilson, Principal of the Gordon Hauck Center; Ruth Hayes, Special Education Administrator; and Joyce Vanden Hoorn, Administrative Assistant. Fred Row, Director of Special Education, and Anne Boone, Principal of the C. O. Sorenson School, represent Northshore. Abby Adams, Director of Special

Education, represents Issaquah. Sharon Hill, Director of Special Education, and Jerry Litzenberger, Director of Research and Evaluation, represent Bellevue. Henry Bertness, Assistant Superintendent; Genevieve Fisher Frankenberg, Child Find/Staff Development Coordinator; and Richard King, Special Education Coordinator, represent Tacoma. Joining the representatives of the cooperating districts is Greg Kirsch, Director of Special Education in the Office of the Superintendent of Public Instruction for Washington State, Division of Special Services and Professional Programs.

Research is also conducted in the Experimental Education Unit (EEU) of the Child Development and Mental Retardation Center, located in the University Affiliated Facility at the University of Washington. Kevin Cole is the principal of the EEU, which serves severely handicapped pupils from several local school districts. Pupils are placed at the EEU when administrators, parents, and teachers determine they would be better served at the EEU rather than in their local programs. All of the pupils live with parents, guardians, or in group homes away from school. The EEU is in session all twelve months of the year, which permits UWRO to conduct research during the summer.

Evaluation

Each of the objectives is a necessary step in achieving our mission. Evaluation of our progress in meeting the objectives is an ongoing process and constitutes Task 3 of the Institute. Three general classifications of data will be collected for overall Institute evaluation: intervention, formative, and summative.

Intervention data, which will serve as the primary source of evaluation, are measures of the actual performance of the subjects during instruction and in nontraining settings, collected during the research activities. UWRO is conducting a wide range of carefully designed studies in an attempt to identify and develop procedures which will help severely handicapped persons to generalize and effectively use the skills they learn. The data collected on the performance of severely handicapped pupils during the research studies will be analyzed by a variety of procedures, including visual inspection of graphed data, trend analysis, and time series analyses for repeated measurement of single subjects. Analysis of group data will utilize correlational and standard tests of statistical significance. Standard analytic practices will determine if an intervention has an effect on subject performance, and the extent of such effects. Data will also be compared to the performance of students participating in the longitudinal study of educational environments, described in Chapter 3.

Some studies will begin in very special settings where the greatest control over conditions can be exerted. It should be noted, however, that all research studies include specific time lines for moving into applied settings—schools, homes, and the community—and evaluating the impact of findings in the "real world." Each line of study is designed to culminate in a material product, such as a manual or set of materials, which describes exactly how the parent, teacher, or other practitioner can use UWRO's findings to facili-

tate skill generalization. Since things which are possible are not necessarily easy or efficient, cost studies will be undertaken to evaluate the time, energy, and resources required to undertake the procedures recommended by UWRO. If necessary, recommended procedures will be modified and retested to make them more easily understood and implemented within the typical applied setting.

Some research designs disregard the fact that statistically significant effects obtained in controlled studies may not have any practical value in normal situations. Therefore, the evaluation of UWRO will rest very heavily on demonstrating that its work is actually having an impact on the lives of severely handicapped persons, not in special laboratories or experimental programs, but in their regular classrooms, homes, and places of work.

The purpose of formative evaluation is to demonstrate the extent that research and communication activities contribute to the development of UWRO's research and attainment of our overall mission. Formative evaluation will incorporate data and descriptions documenting existing research and demonstration procedures reviewed, applications and adaptations of existing procedures, and research activities utilizing existing procedures. Products which result from UWRO's contacts with other Institutes and agencies will be reported. Any products disseminated as a result of either individual studies within UWRO or through contact and collaboration with researchers outside UWRO will also be documented.

Summative evaluation contributes to assessment of the lasting impact of UWRO's various activities. We will collect and analyze data on (1) the extent to which UWRO's research and intervention procedures and materials are adopted by local educational agencies, (2) evidence of the quality of research skills acquired by UWRO Research Assistants, (3) evidence of cost savings resulting from implementation of UWRO-developed procedures, (4) the adoption of UWRO-developed procedures in teacher preparation and inservice training courses, (5) changes in peer interactions resulting from UWRO research and intervention procedures and products, and (6) the overall attainment of UWRO's goals. These data will be available to the other Institutes and researchers in the field, for their information and to stimulate interaction with UWRO. This documentation also will serve as evidence of any "ripple effects" produced by UWRO efforts.

Finally, the overall impact of UWRO's efforts must be measured in terms of how much generalization is facilitated. This will be evaluated by conducting extensive inservice training seminars for teachers, parents, administrators, and other training professionals and determining, through follow-up evaluation, the extent to which the participants adopted the procedures and whether there was any demonstrable beneficial impact on the lives of their children or clients by increasing the nature or extent of generalization from "pre-UWRO" levels. That will represent the most meaningful evaluation of the UWRO.

Communication

Communication about ongoing studies, training, and dissemination of procedures and products derived from our research is an important component

of the UWRO mission. Cooperation with other researchers exploring issues related to skill generalization is the necessary first step in maximizing the potential benefits of UWRO activities. We will be working closely with the three other Institutes for Research in Education of the Severely Handicapped.

Two of these Institutes will focus on methods of facilitating integration from restrictive to least restrictive educational environments. The University of Minnesota's "Consortium Institute for Education of Severely Handicapped Children" directed by John Rynders, and San Francisco State University's "California Research Institute on Transition of Severely Handicapped Students to the Least Restrictive Environment" directed by Wayne Sailor, will study the ways and means of integrating severely handicapped students with their nonhandicapped peers. Another Institute will also study generalization. "Extending Competent Performance: An Institute for the Study of Generalization with Severely Handicapped Students" is under the direction of Robert Horner at the University of Oregon. Since the process of integration is likely to involve the necessity of generalized responding in "new," integrated environments, the work of each Institute will relate directly to the work of the others.

The four Institutes are committed to maintaining active interaction with one another. Researchers may assist each other by replicating various procedures or interventions. Conclusions drawn at one Institute may be incorporated into designs for studies at other Institutes. We will also be able to share our failures—important information that is seldom published—to prevent investigation of ineffective procedures. Methodological problems and solutions can be shared, preventing duplication of mistakes. Under normal circumstances, new data are seen by other researchers only after they appear in a professional journal. Since the publication process often takes as long as two years, relevant data may not be available when needed. This problem will be circumvented by monthly communication and inter-Institute meetings. Data from the other three Institutes will affect the direction and content of our research, and stimulate creative approaches to our work.

The second important communication activity is training. This activity will commence with the training of Research Assistants by Senior Investigators. Efforts will be made to employ Research Assistants who are students enrolled in graduate programs in Special Education and related fields. These potential researchers and professionals will acquire training in those skills required to conduct different types of research in laboratory and applied settings. At the same time, they will acquire experience in promoting generalization in educational settings. We expect training of Research Assistants to be conducted continuously during the project.

The results of individual studies in generalization conducted by UWRO, as well as the results of studies conducted within other research Institutes, will be disseminated via courses taught by Senior Investigators who are also teaching faculty at the University of Washington. Consultations and workshops given by Senior Investigators at other agencies, colleges, and universities may include results to date.

During the third project year training in instructional and curricular procedures will involve personnel from Direct Service Consortium schools. Training will be conducted by UWRO staff. The nature of the training will depend on the requirements of the local educational agency and will focus directly on the application of procedures in natural settings. Opportunities for training will be extended to personnel from all local educational agencies in Washington during the fourth project year in cooperation with the Office of the Superintendent of Public Instruction for Washington State. As information from the "guideline" studies is collected, training content will be modified. During the fourth project year, it is expected that the training will emphasize the guidelines for each area and practical methods of integrating the approaches in educational settings.

Training is perhaps the most active communication process, but it will reach only a small percentage of interested professionals. In order to increase the potential benefits of UWRO procedures, technical information and the guidelines for practical application will be disseminated through publications and direct mailings.

Technical information will include precise and detailed descriptions of research methodology, analytic procedures, the relationship between the research conducted by UWRO and the existing body of research information, and presentation and discussion of the results of individual studies. Technical information will be in the form of individual articles prepared for journals, in proceedings from the inter-Institute conferences, in annual "Review of the Literature" publications, in Annual Reports from UWRO, in the Final Report, and in the Research Monograph to be produced during the fourth and fifth project years. This information will also be disseminated through discussions with researchers at inter-Institute meetings, at national conferences, and at a series of quarterly professional seminars conducted at the University of Washington.

UWRO will produce several publications of "best practices" guidelines. These materials will be assembled for specific audiences, including teachers, teacher trainers, parents, supervisors, administrators, curriculum specialists, and related professionals. A wide variety of persons interested in the research will receive this practical information, which will be disseminated via training, presentations at conferences, and mailing of project products. National dissemination targets will be identified, but persons interested in receiving project information will be able to contact UWRO directly and obtain any product at a small cost.

UWRO activities will be of little ultimate value if the results are not available to those who need them. Communication activities will include cooperation, training, and product dissemination.

Advisory Committee

To ensure that research will have practical application to a wide variety of potential consumers and to provide advice from professional perspectives,

administrators, parents, researchers and others met during the formulation of the UWRO proposal. Now meeting as the Advisory Committee, they provide advice on ongoing activities and assist the project in maintaining a practical approach to the education of the severely handicapped. The members represent the full range of professional activities and service delivery systems in the State of Washington.

Local educational agencies are represented by two individuals: Genevieve Frankenberg, Coordinator of Child Find and Staff Development for Tacoma School District No. 10, is the Advisory Committee's Chairperson and is also a member of the UWRO Direct Service Consortium. Fred Row, Director of Special Education for Northshore School District No. 417, is also a Direct Service Consortium member. Intermediate educational agencies are represented by Donald Whitney, Director of Special Services for Educational Service District No. 121, which serves 36 local education agencies. Judy Schrag is the Assistant Superintendent for the Division of Special Services in the Washington State Office of Superintendent of Public Instruction and will be our liaison with the state educational agency. Al Bauer represents the 49th Legislative District in the Washington State Senate and sits on the Rules, Ways and Means, Financial Institutions, and Local Government Committees. Joseph Jenkins is Director of the Experimental Education Unit and a member of the faculty of the College of Education of the University of Washington. A noted researcher, he will contribute information from the perspective of a researcher and as an representative of an institution of higher learning. Margo Thornley is Executive Director of the Wiser Vocational Institute, which provides vocational evaluation and training to severely handicapped individuals. She represents other service agencies on the Advisory Committee. Kathleen Knowlan is a student in Speech and Hearing Sciences at the University of Washington and is completing a BA in Communication Disorders. She plans to complete a graduate program in Clinical Speech Pathology. She is the parent of a handicapped child. Together, these individuals bring a wide background of experience, a variety of perspectives, and a sincere interest in the education of the severely handicapped to assist the Washington Research Organization in meeting its goals.

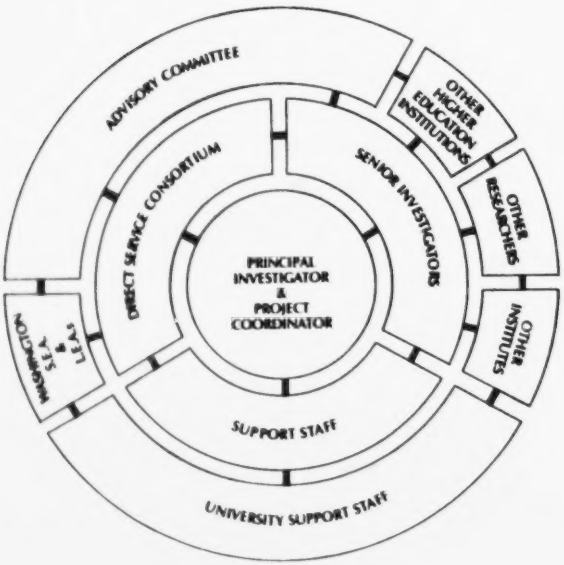
Administration and Management

Administrative activities support the research, evaluation, and communication tasks of the Institute. General administrative tasks relating to employment, personnel management, purchasing, budgeting, and federal reporting requirements are covered by this task.

While such administrative tasks are conducted in every organization, the structure of UWRO is designed to facilitate our unique activities. Rather than an hierarchical system where responsibility and information flows from "top to bottom," we have a circular structure. Information flow is both circular, within the rings, and linear, to and from each circle. In addition, most personnel will participate in more than one group, thus increasing the nonhierarchical structure of communication.

The overall responsibility for UWRO activities rests with the Principal Investigator and the Project Coordinator, but decision making is shared by all groups. Individuals will make decisions related to activities for which they are responsible. Decisions for group and intergroup activities will be reached by consensus. The model for communication at UWRO is shown in the following chart.

UWRO Organizational Structure Chart



TWO

BASIC RESEARCH CONCEPTS

Skill Generalization

Sometimes we want generalization to occur and other times we do not. For example, if we are successful in decelerating or eliminating maladaptive behaviors during training, such as spitting and hitting, we want those behaviors to not occur in other environments. The aim of such programs is generalization of nonresponding. Since special conditions and circumstances surround this kind of training, and since instances of maladaptive behaviors may actually decrease as skill and competence increase, very little of UWRO's research will be concerned with the generalization of "no response." Most of the time, as educators, we do want generalization to occur. If we train toileting at home, we want to see toileting at school. Generally, the behaviors we train may be called "skills" or "skilled behaviors," because they provide the student with the competencies needed for normal living; these will be the ones of most interest in our investigations.

Broadly speaking, skill generalization is appropriate responding in the absence of programmed training procedures. Severely handicapped individuals are taught specific responses under special conditions involving instructional techniques developed through experimentation. These techniques involve a variety of elements, including the events that immediately precede the response, such as verbal directions (e.g., "Get dressed." "Put on your shoes."). These antecedents may come to control the response and are then called discriminative stimuli for responding. The student responds when discriminative stimuli are present, and does not respond when they are not. Other discriminative stimuli may include specific materials or objects (e.g., T-shirt, shoes) used during instruction, the setting of the instruction (e.g., the desk, the room) and the trainers involved. Instructional techniques also involve events that follow one or more responses, like praise or candy (e.g., "Yes, that's the way to get dressed."), or feedback on incorrect responses (e.g., "No, that goes on your other foot."). These events are called consequences. Consequences are usually arranged to follow the response; their occurrence is contingent upon the response. Contingencies are programmed during training, and may vary from one consequence for each response to one consequence for several responses.

Generalization is concerned with the performance of the response outside of training settings. When the specific events that occurred during training are not available, different stimuli may serve to signal the response. Outside the training setting, contingencies for responding are different; consequences may or may not follow the response. In analyzing why generalization does or does not occur, investigators have found it useful to examine separately each of the areas where differences exist: in the stimuli, people, consequences, settings, and over time.

When the student responds appropriately to untrained instances, objects, or cues, "generalization across stimuli" is said to occur. For example, instruction in putting on shoes may have included only loafers; if the student is able to put on a slipper, using the same motor skills in her response, generalization across stimuli has occurred. In cases where generalization does not occur, it is hypothesized that discrimination training has been so successful that the student will respond only to stimuli that are identical to the training stimuli. When the stimuli change, the student "recognizes" the change, and thus does not respond. If the student does respond to stimuli that are similar to the trained stimuli, then generalization has occurred.

Another problem area in generalization appears to involve the trainers. Often trained responses occur only in the presence of the people who trained the response, even if the same antecedents and consequences are involved. When the student responds appropriately to people who did not train him, "generalization across people" is said to occur. For example, if the student has been taught to say, "Hi, my name is Charles," and is able to respond to a stranger's introduction with those appropriate words, generalization across people has occurred.

Many instructional situations, especially during skill acquisition, involve consequences for each response. One to one contingencies are unusual outside of acquisition programs. Also, the consequences available during instruction, such as candy or hugs, may not be as available after instruction ceases. When the student responds appropriately in the absence of the consequences available in the training environment or to different contingencies of consequation, "generalization to natural consequences" is said to occur. For example, training procedures may have included candy for each correct response. If the student responds appropriately and continues to respond with only intermittent praise, generalization to natural consequences has occurred.

"Generalization across settings" is a broad descriptor which incorporates each of the types described above and generally defines the incredible variety of changes that occur when the student is expected to respond in new settings. For example, training a pupil to identify buses by number, to enter the bus, to pay for his fare, and to exit from a bus at his destination may all occur within a classroom setting. However, the student must be able to apply this learning to actual travel. If successful, generalization across settings has occurred. The differences between the training setting and the actual use of city buses are so many and so varied that this category is used to describe the collective differences.

We include another category of generalization, "generalization across time." If the response continues to be performed appropriately after training ceases, generalization across time has occurred. This is also called "maintenance" or "retention," but since training has ceased, the conditions have changed (i.e., antecedents and consequences may be different or presented irregularly), and thus may also be appropriately classified as generalization.

So far, generalization has been described as occurring when the trained response is performed under untrained situations. However, the true pur-

pose of teaching generalized responding is to provide the individual with means of adapting to new situations, solving problems, and living in different settings. The response must be appropriate. "Hi, my name is Charles," may be said perfectly in a new setting, but if it follows the stimulus, "Put on your jacket," it is entirely wrong.

If the true aim is getting along in new environments, then the response must also be modifiable, or physically adapted, to the setting. Many instances of generalization involve changes in the physical actions that constitute the response. For example, putting on a T-shirt with long sleeves requires slightly different physical movements than putting on a short-sleeved T-shirt. In other cases very different physical responses will be required to achieve the same effect as that achieved by performing the trained response. For example, training a student to put on a shoe achieves the effect of covering and protecting the feet. Putting on a pair of rubber boots achieves the same effect, but physically different responses are usually involved.

Other problems must be solved if the student is to respond successfully in new environments. One method of solving problems in new environments is to combine two or more responses that were learned separately. For example, a student may learn how to reach for and grasp something on a shelf above her head. In another training situation, she may be taught how to stand on a chair. If she were to successfully use both of those skills to get her lunch from a high closet shelf without training or prompting, she would have solved a typical problem situation that may occur whenever she is in a different environment. When decisions are required, a response adapted, or a problem solved, generalization involves much more than simple adaptation of a learned skill; it involves adaptation. UWRO investigators will study both types of generalization: application and adaptation.

Instructional Programming for Generalization

Until recently, many people expected generalization to occur spontaneously after training; a "passive" approach to instructing for generalization has been common. We know that the "train and hope" method does not result in much generalized responding by the majority of severely handicapped students. Trevor Stokes, of the University of Manitoba, and Don Baer, of the University of Kansas, published a major analysis and summary of research in generalization in 1977. This article, and the discussion it provoked, had a major impact on shaping subsequent research in generalization. They argued that it is better to view generalization as an active process and to try to develop instructional methods that ensure that generalization does occur.

Stokes and Baer identified methods in addition to "Train and Hope" that had been reported in published research. In "sequential modification," the behavior is trained in one setting and then, if generalization does not occur in the next setting, training is programmed for that setting, and so on for each setting. This is actually not a practical solution to the problem by itself, since it would mean that training would have to occur in every setting and each time the individual moved to a new setting.

A similar technique, which Stokes and Baer feel is more promising, requires introducing many different types of similar antecedents into the training situation. By "training sufficient exemplars," the individual is thought to learn a general category of items or objects to which to respond. For example, instead of teaching "putting on a sweater" with just long-sleeved crew-necked sweaters, V-necked sweaters, short-sleeved sweaters and so on are trained. With more varied instructional antecedents, generalization to untrained sweaters (e.g., turtlenecks) may occur.

An extension of this technique was classified by Stokes and Baer as "train loosely," in which many different antecedent events are introduced during training. For example, instead of prefacing each trial with, "Put on your shoes," the student may hear, "Put it on," or "It's time to go outside, shoes on," or even be given the shoes without any verbal direction. The more specific "program common stimuli" technique would be to identify common elements in different environments and include those in the training setting.

In addition to problems associated with antecedent stimuli, it has been hypothesized that generalized responding does not occur or maintain because the consequences available in natural settings either are not reinforcing to the individual or do not occur as frequently as they did during training. Research data has already shown us that if a frequently reinforced response is performed under infrequent reinforcement, that response is likely to disappear—to be extinguished. The technique "use indiscriminable contingencies" involves gradually replacing training consequences and schedules with the contingencies of natural settings. In this manner, naturally available consequences acquire reinforcing powers through pairing with programmed consequences, before training consequences are discontinued. Similarly, the schedule of one consequence for each response commonly used during training is gradually replaced with a schedule of intermittent consequences, so that the student is unable to discriminate when a response is likely to be reinforced and when it is not. Thus this method is designed to ensure that generalized responding will occur and endure with the infrequent natural consequences available outside of training settings.

Another method identified by Stokes and Baer is to gradually introduce the individual to "natural maintaining contingencies." This can be achieved most easily by teaching behaviors that are functional in nontraining settings. For example, teaching appropriate eating behavior would introduce the child to the contingencies that occur naturally, such as compliments, access to different foods, opportunities to eat at restaurants, or outings with family and friends. The natural consequences would then reinforce continued "good eating behavior." The student is introduced to natural consequences by teaching her a response that will be naturally reinforced in normal settings.

Another technique that has been used to "train to generalize" seems to be at odds with most established instructional methods directed at acquisition. In this method, consequation occurs only for generalized responding. In such situations, the learner would not be reinforced for learning a new skill, but only for using it appropriately outside of training situations.

A final category of research involves processes that Stokes and Baer identified as "mediate generalization." Teaching the individual new methods of thinking and acting or to use self-control strategies are examples of teaching "mediated generalization" skills, rather than directing programming at generalization of specific skilled responses.

Each method shows some promise but to date no approach has demonstrated consistently good effects in controlled settings, and little research has been conducted in classrooms and homes with teachers and parents implementing the procedures. Our research will seek to extend and develop these and other approaches to the problem, using the methodology discussed in the next section.

Methodology

Subjects and Settings

The subjects will be students attending the Experimental Education Unit and schools of the Direct Service Consortium who meet local, state, and federal classifications as severely handicapped, profoundly handicapped, severely behaviorally disordered, autistic, childhood schizophrenic, deaf-blind, or multiply handicapped. In order to facilitate the identification of subjects while respecting the Rights of Human Subjects Guidelines, these districts will write letters to parents of students explaining the research. Parents or guardians will be given the opportunity to voluntarily consent to their child's participation in a specific study. Teachers of students for whom consent is obtained may also consent to participate in research studies.

Since we are investigating skill generalization, measurement of generalization will occur in a wide variety of natural educational settings, including classroom, school, home, community, and vocational environments. Some studies will involve subjects working directly with a researcher in a separate room or in a part of the training or classroom setting. Results of research in such controlled settings will be applied to more normal settings. Other studies will involve students working individually or in a group with their regular classroom teacher. Studies in nonschool environments will involve parents, supervisors, peers, neighbors, or others in the subjects' normal daily routines.

Subject Responses

In each study, the performance of the student will be measured. Such measurements are used to determine the effects of different types of training, the effects of changing trainers, the effects of changing settings. Performance data will provide the information we need to better understand the phenomenon of generalization and practical methods of achieving it for many different individuals and many different skills.

The selection of skills or behaviors to be measured will be determined by methodological factors, but will also involve educational considerations of functionality and age-appropriateness. Increasing emphasis is being placed on teaching severely handicapped individuals functional and age-appropriate skills. It may be appropriate to teach playing with blocks to a preschool child as a leisure skill, but bowling is a far more age-appropriate leisure skill for a teenager. This concept also extends to the selection of instructional materials. While beads and blocks may be appropriate materials to teach a youngster to discriminate objects by shape, spoons and forks are more appropriate for teaching the same skill to a teenager. Furthermore, rather than teaching skills with limited use in most daily environments (i.e., making holiday ornaments), teachers are now concentrating instruction in areas more relevant to daily living and vocational success and ones which introduce the pupil to natural maintaining contingencies.

Some of our research will involve collecting data on behaviors targeted on IEPs, but other factors must be considered. For example, generalization might better be achieved for the group of dressing skills if training included practice with a wide variety of different types of clothing items (e.g., sweaters trained: cardigans, pullovers, zipback sweaters, v-necked sweaters, etc.), than with repeated practice on items of a single type. Conversely, generalization may be hindered for the class of grooming behaviors by training in a wide variety of items, but facilitated by teaching the student to check his own appearance. In some studies, therefore, grooming might be selected, while in others, dressing. As information about generalization accumulates, factors such as these may influence the selection of student responses for study.

In the selection of subject responses it is necessary, especially in the early stages of research, to make sure that any observed changes are the result of the intervention being tested. The researcher may need to ensure that experimental instruction is the only training affecting the performance of the subject. This control is very difficult to achieve if common functional skills are selected. How can the researcher who selects dressing skills be sure that instruction is not being conducted in the home or school, even incidentally? In order to eliminate such effects, tasks which are relevant only to the study may be selected. Effective strategies identified by studies measuring artificial experimental responses will be applied to functional tasks later on.

Measurement of Generalization

The basic concerns in collecting data on generalization include, in addition to standard research concerns of reliability and validity, the scheduling of generalization "probes" or measures, the frequency of measurement, and the quality of the generalized response. In most published research, generalization is measured by one or more "probes" or "tests" following the conclusion of training or after the subject has met a predefined criterion performance level on the trained skill. These data can provide us with evidence that generalization did or did not occur. However, if measured "after" only, we really can't determine when generalization began or compare performance with preintervention levels.

Measuring generalization both before and after provides information on the net impact of the training, but leaves other important questions unanswered. Does generalization begin to occur gradually, paralleling acquisition of the skill, or only as some level of mastery is reached? Do different methods of teaching generalized responding promote generalization earlier than others? Does generalization occur soon after training begins or only toward the end of training? With such information, we can begin to understand the relationship between skill acquisition, fluency-building, application, and adaptation. These questions can only be answered by measuring generalization during training and repeatedly over time, as UWRO will do. Repeated measures or opportunities to perform the generalized response will provide information on the progress of generalization as an ongoing, active process, rather than as an single spontaneous event. Not only will data be collected at different times in relation to an intervention, but multiple probes will be scheduled at each time.

Repeated measures of generalization will also provide information on another aspect of generalization, one that has received little attention: "training savings." An individual who been successfully instructed in one skill may learn another skill very quickly as a result of the previous instruction. For example, a student may complete all of the steps required to boil an egg accurately (i.e., without breakage or overcooking) and fluently (i.e., in the time it takes an average adult to boil an egg) in seven training sessions of 15 minutes each. Following egg training, the student may need only one session to master broccoli cookery. This may be compared with another student who was taught to boil broccoli without egg training, and who took eight sessions to achieve the level of mastery the first student achieved in one. This "savings" of time spent in instruction is another important dimension in building generalization skills and is of practical significance to educators.

Measures of generalization often include only "yes/no" data on whether generalization occurs and/or a statement of the accuracy of performance (e.g., 80e "time" of the response, expressed as either rate, latency, or duration, is required in addition to accuracy data in order to understand how severely handicapped individuals acquire and build fluency in skills. The length of "waiting" time before responding (i.e., latency), the rate of responding, and the duration of the response itself each provide important information on the quality of the response.

A good example of the importance of the temporal quality of generalization is dressing. If a child is taught to dress herself accurately and to finish within 10 minutes during training, it is important to know not only whether or not dressing occurs at home (i.e., yes/no data), how many items of clothing are put on correctly (i.e., accuracy data), but also how long it takes her (i.e., fluency data). Presumably the 10-minute training criterion is set to allow the child to complete dressing within a time limit that is functional for her home environment (e.g., morning schedule does not allow for more than 10 minutes for dressing). Thus, if training in dressing produces accurate and speedy dressing in the home, the training may be regarded as entirely successful. However, what is the quality of generalization if the child dresses accurately, but takes 35 minutes to do so? Obviously, the significance of the generalization achieved is less than in the former case. Thirty-five minute dressing may

even have more serious consequences for behavior maintenance. The parents, anxious to see the child dressed and breakfasted before the school bus arrives, may decide to "help" the child dress or even dress her themselves. Over time, the opportunity to dress is withdrawn and we would expect that the skill of dressing may be lost. In order to measure all of the important dimensions in generalized responding, UWRO researchers will collect yes/no, accuracy, and fluency data as measures of skill generalization.

Procedures for Descriptive Studies

UWRO's research activities will begin with studies designed to provide additional information about variables already identified, such as the stimuli, contingencies, consequences, settings, and conditions in environments where generalized responding is desirable. Descriptive studies will also include examination of other variables that may affect skill generalization, such as the scheduling of instruction or the learning characteristics of the individual.

The collection of descriptive data will involve three different types of analyses. Analyses of data collected previously may be used to generate hypotheses, since it is unlikely that the experimenter's bias could affect the data. Similarly, an analysis of published research using statistical summaries across studies and discriminate analysis techniques may provide additional information. Descriptive studies will also include actively collecting data in educational settings, without any intervention.

Procedures for Intervention Studies

Intervention studies include both controlled laboratory studies and investigations in applied settings. The intervention research will utilize two distinct methodological approaches in investigating generalization in severely handicapped individuals: "single subject" and "group" designs. In each methodology, our interest is in determining the effects of various interventions on skill generalization or response adaptation of the subjects involved in the study.

Single subject research designs include repeated measurement of the target behavior, and thus provide information on the process of change of the behavior. Data are collected on the target behavior over a period of time before an intervention is introduced. The effect of the intervention is determined by comparing performance before, during, and after the intervention. The relative strength of an intervention is tested by withdrawing the intervention and analyzing any changes. If the intervention cannot be withdrawn, as when an intervention has taught a new way of responding, the intervention is implemented with other behaviors and with other subjects. The data collected on each subject are studied individually and analyzed to determine the process of change involved. Replicating the studies will provide information on the generality of the results.

In group designs, subjects are selected to be representative of a large population and then randomly divided into two or more groups. Sometimes a single measure of performance of the target behavior is used as a pre-measure or pre-test. One group is chosen as the control group and another as the intervention group. There may be several different types of interventions tested, but usually only one per experimental group. Following the intervention, a post-test or measure of performance is taken. The effects of the intervention are determined by comparing the performance of the experimental group with that of the control group. The data on each group are studied as a single unit to determine the product or net effect of the intervention. The performance of a single individual is important only as an indication of individual differences within the group. Inferences and results obtained by studying a group may lead to information about how procedures may be likely to affect the population from which the group was originally drawn.

THREE

UWRO'S APPROACHES TO GENERALIZATION

UWRO's research will involve four different but interrelated lines of inquiry to approach the fundamental questions about generalization: Why do some students generalize and others not? What can we educators do to see that all students are able to generalize? These approaches are distinguished by their basic assumptions and by the types of the intervention strategies investigated. The four approaches to these questions are:

(1) an "ecological" approach to describing and then changing conditions within the educational environment;
(2) a "performance pattern" approach to describing and matching individual learning characteristics and instructional techniques;
(3) a "self-control" approach to teaching severely handicapped individuals to manage their own behavior; and
(4) "strategy implementation" investigations to test the effects of strategies identified by the other three approaches when they are implemented in typical classroom settings.

The time lines for UWRO research activities proceed generally from descriptive studies and tightly controlled laboratory intervention studies to intervention studies in natural environments. The longitudinal descriptive study of existing conditions in training and nontraining settings will continue throughout the project. The hypotheses of the studies designed to intervene in existing ecological conditions are, of all of the areas, most firmly rooted in existing research. Therefore, intervention studies will begin initially in applied settings (e.g., public school classrooms).

The performance pattern research will begin with descriptive studies involving retrospective analyses of existing data sets, and then proceed to the collection of descriptive data in public school classrooms. The descriptive information will be used to determine a set of experimental data decision rules for matching specific instructional methods to individual performance which will be tested in intervention studies in applied settings during the third and fourth project years.

Without existing data sets or even very much applicable literature, studies in self-control will begin with tightly controlled intervention studies under laboratory conditions. Each self-control skill will be investigated in the laboratory before intervention research begins in applied settings.

Beginning in the third year, strategy implementation studies will investigate strategies developed by researchers in the ecological, performance pattern, and self-control studies. These studies will test the effects of UWRO's generalization strategies when they are implemented by public school teachers under typical classroom conditions.

RESEARCH TIME LINES

	Ecology of Training Settings	Performance Pattern Studies	Self-Control Studies	Strategy Implementation Studies
Year 1 82-83	**Descriptive** Environment X Indiv. Characteristics (1st group) **Intervention** (Applied) Massed v. Distributed Trials	**Descriptive** Existing Data Sets	**Intervention** (Laboratory) Self-Monitoring	
Year 2 83-84	**Descriptive** Env. X Charac. (1st & 2nd group) **Intervention** (Applied) Competing Behaviors Instructional Schedules Environmental Simulation	**Descriptive** Existing Instruction X Perf. Patterns	**Intervention** (Applied) Self-monitoring **Intervention** (Laboratory) Self-Reinforcement	
Year 3 84-85	**Descriptive** Env. X Charac. (1st, 2nd & 3rd Groups) **Intervention** (Applied) Instructional Schedules Environmental Simulation	**Descriptive** Existing Inst. X Perf. Patterns **Intervention** (Applied) Data Decision Rules	**Intervention** (Applied) Self-Reinforcement Self-Monitoring	**Intervention** (Applied) Ecological Strategies Self-Control Strategies Perf. Pattern Strategies
Year 4 85-86	**Descriptive** Env. X Charac. (1st, 2nd, 3rd & 4th Groups) **Intervention** (Applied) Environmental Simulation Guidelines	**Intervention** (Applied) Data Decision Rules Guidelines	**Intervention** (Laboratory) Self-Instruction Combined Strategies **Intervention** (Applied) Combined Strategies Guidelines	**Intervention** (Applied) Ecological Strategies Self-Control Strategies Perf. Pattern Strategies

According to the contract plan, the research activities of the Institute will be concluded by the first part of the fifth project year. Our final activities will emphasize dissemination of the research findings. The background, design, expected results, and findings to date of each of these approaches will be discussed in this chapter.

Studies in Ecological Variables

While educators frequently lament the failure of pupils with severe learning handicaps to generalize, research has suggested a variety of instructional strategies which could potentially be applied in educational programs to increase the probability of obtaining generalization. However, the degree to which these practices have been incorporated into education is unknown, as are factors in educational settings that may limit the effectiveness of these strategies. We will use the word "ecology" to refer to the total of all observable factors and conditions which comprise the educational setting. The purpose of research in this area is to explore the current educational ecology of severely handicapped pupils and selected ways in which ecological conditions might be modified to enhance generalization.

Ecology studies will begin with a four-year descriptive study. The longitudinal descriptive study will serve several purposes. The data collected will be used as a general baseline for all UWRO research, to determine the extent to which our procedures are adopted, as a basis for cost comparisons, and as a general indication of the level of generalized responding with and without UWRO procedures. However, the primary purpose is to explore a variety of the factors in educational settings that may influence generalization. Factors so identified will be investigated in a series of intervention studies.

Design of Longitudinal Descriptive Study

This four-year study will explore four major issues:

(1) The nature and number of pupil goals and objectives that include the intent to promote generalization or include behaviors that require generalization in order to be of functional value.
(2) The extent that pupil performance data indicate attainment of goals/objectives related to generalization.
(3) The degree to which generalization occurs as a result of formal or informal programming.
(4) General ecological conditions which might facilitate generalization.

Ecological conditions which may be examined include the number of managers administering formal or informal programs throughout the day, the number of intraschool environments in which programming occurs, the percent of the school day in community environments, the degree of interaction and opportunity for interaction with nonhandicapped or with lesser

handicapped peers, and the number of school/community cooperative programs administered following school hours (e.g., programs administered by parents). Other conditions may be identified as the study progresses.

The descriptive study will include two types of activities. First, a review of existing records (e.g., IEPs, formative data, and lesson plans) for a sample of severely handicapped pupils selected from Direct Service Consortium schools will be undertaken to collect information on objectives and educational plans. Second, interviews with teachers, parents, and/or other caregivers, as well as data collection in classrooms and other environments by members of the project staff, will gather information related to current conditions, implementation of procedures, and pupil performance.

Selected members of the original pupil sample will be followed each year for three additional years to develop a longitudinal record of generalized skills as they are acquired. This study will also record the changing nature of conditions to which pupils are exposed over time. In addition, new pupils will be selected and added to the sample each year, with similar information collected on each.

Analytic techniques applied to the data will be primarily descriptive and exploratory in nature. Ultimately, the results will be of value in determining which currently employed instructional strategies have a high probability of success, whether certain skills are more likely than others to generalize in the absence of formal programming for generalization, and the basic conditions within community, home, and work environments which should be considered when designing, implementing, and evaluating programs to facilitate generalization for severely handicapped persons. The data collected will be used in the selection of ecological intervention studies and in the selection of variables and methodologies for other UWRO investigations.

Design of Intervention Studies

The second set of ecological investigations will study interactions between factors identified in the descriptive study and the effect of changing one or more of the existing conditions on the occurrence, quality, and quantity of generalized responding. Factors that are likely to be studied include pupil response variables, general task and setting variables, stimulus variables, response demands, and reinforcement variables as they pertain to the development of generalization.

Although it is impossible to predict all factors that will be investigated, previous studies have identified likely areas. Past research leads us to believe that when generalization does not occur, the individual may have previously acquired a behavior that achieves the same outcome as the behavior you are training. This behavior may compete with the trained behavior in the critical effect of a response. For example, tantrumming to be fed may compete with feeding oneself or asking for food, if each gets the same results. At least one study will be conducted to investigate methods of

identifying and managing undesirable competing behaviors and assessing the impact of the interventions on generalization of more desirable responses.

Another intervention study will examine the relationship of generalization to the scheduling of instructional trials. Instructional opportunities, or trials, are frequently grouped into a single block, with one trial immediately following the next. For example, 10 or 15 trials of "buttoning" instruction might be presented daily from 10:00 to 10:30 a.m. An alternative to this practice would be to provide buttoning instruction at times when a natural need exists to button one's clothing (e.g., upon getting up in the morning, before going outside, after gym, or after using the toilet). This method for scheduling instruction would result in the provision of trials spaced or naturally distributed throughout the day.

Scheduling instruction at the times when the target behavior would occur in natural environments could prove beneficial for several reasons: (1) It would increase the similarity between the instructional situation and the conditions in which generalized responding is desired. (2) It might increase the likelihood that unprogrammed reinforcers would be available in the natural, generalized setting. (3) It may avoid problems often noted with the severely handicapped, like "poor attention spans," fatigue, and reinforcer satiation. This series of investigations will provide data on the relative impact of different trial scheduling formats on generalization.

The methodology of the intervention studies will be single subject designs replicated across subjects. Repeated measurements of pupil performance during the instructional sessions and in the generalization setting(s) will serve as the primary independent variable. In cases where training occurs in the "natural environment," generalization will be measured in different but similar settings. Analytic techniques will include visual inspection of graphed data, time-series statistical techniques, and an overall statistical summary of performance for comparison between studies.

Expected Outcomes and Products

The studies in this area should result in the development of a "best practices" manual, incorporating at least four areas:

(1) guidelines which identify current best practices existing in public school settings will result from the descriptive studies,
(2) guidelines for identifying competing behaviors and methods for counteracting their effects in nontraining environments,
(3) guidelines for how to schedule instructional trials and learning opportunities for different classes of skills, and
(4) other possible guidelines may be developed depending on information from the descriptive studies and the nature and results of the intervention studies.

Summary of First Year's Findings

Five descriptive studies of the ecological conditions of educational settings were initiated during the first project year. Two to four research assistants

worked with Dr. Felix Billingsley on the longitudinal ecological studies, which were conducted with subjects from two local public school districts. For more information on these studies, see Billingsley, Berman, & Opalski (1983).

The general approach of this investigation was to measure, through direct observation, surveys, interviews, and/or record reviews, selected variables within the educational ecology of severely handicapped pupils. Those measurements were then employed to determine the match between existing conditions and those conditions which ought to exist if the principles outlined by Stokes and Baer (1977) were generally applied.

Study I: Generalization in IEPs. The first study was undertaken to determine whether teachers indicated the intent to promote generalization across situations or persons in IEP objectives, and whether the objectives were of such a nature that skill performance was likely to have adaptive value. IEPs for the 1981-82 and 1982-83 school years (22 IEPs per year) were reviewed.

Perhaps the most dramatic findings were related to the variability in total number of objectives included on IEPs and the very small percentages of objectives which specified generalization intent. The total number of objectives included on IEPs varied from 2 to 26, with a median of about 10.5. For at least some children, the number of skills being taught which could potentially be generalized was quite low.

In both sites, the number of objectives in which generalization intent was specified was negligible. Overall, 7% or less of the coded objectives specified generalization intent. If IEP objectives accurately indicate desired instructional outcomes, then it must be concluded that generalized skill performance was not a high priority with the members of the educational team responsible for writing the IEPs reviewed.

Combining sites, 66% of objectives coded from 1981-82 IEPs were considered functional as were 65% of the objectives from the 1982-83 IEPs. There is no empirical basis upon which to determine whether percentages in the mid-60s are to be commended or criticized. If, however, these data indicate that approximately one-third of the objectives possessed minimal adaptive value and were unlikely to generate reinforcers in nontraining settings, then it may be that teachers are spending considerable time teaching behaviors which have relatively low probability of generalization and, therefore, minimal utility to the learner.

Study II: Parent ratings of objectives and student opportunity to generalize skills to home. The next objectives of the project were to:

1. Determine the opportunities available to utilize instructed skills in home environments, since opportunities to perform are necessary in order for generalization to occur. Furthermore, we would expect instruction to be aimed at skills which the student will in fact have an opportunity to use.
2. Determine if skills were appropriately or inappropriately performed in the home, since generalization could not be said to have occurred if the skill is performed appropriately in the classroom

but not elsewhere. In addition, if a skill is already performed appropriately in the home, the value of further instruction is questionable.

3. Determine if parents were providing training in the home, to see if such training or the lack thereof has an effect on generalized performance.

On the whole, the parents felt that IEP objectives clearly stated the required behaviors. They also indicated that they considered the vast majority (91%) of the behaviors taught to be functional. Parents at both sites indicated that, outside of school, their children had the opportunity to perform only slightly over 60% of the behaviors included in 1981-82 or 1982-83 IEP objectives.

Regarding appropriate performance of behaviors outside of school, parents indicated that approximately 50% of IEP behaviors did occur at appropriate times. In other words, about half of the behaviors targeted for instruction in school were said to be occurring in nonschool settings. This is, of course, good news in the sense that generalized performance of skills seems to be occurring at a level which is likely to have a meaningful impact on the lives of the pupils. It is not such good news, however, in the sense that a number of skills on which instruction was in-progress, appeared to have already been acquired by the subjects.

Relatively few behaviors were noted to occur at inappropriate times within nontraining settings. However, over one-fourth of the parents of subjects at one of the sites indicated that behaviors included in 1981-82 objectives occurred at inappropriate times.

Parents reported home training for 57% of 1981-82 objectives, and for 54% of the 1982-83 objectives. This is an encouraging finding which is generally consistent with principles for development of generalized responding; however, the frequency and precise nature of the home training is unknown.

It was found that 85% of the behaviors which parents indicated were being performed appropriately were also being trained, at least to some extent, at home. This degree of concordance tends to support the possible value of sequential modification and underscores the importance of parent involvement in the educational process.

Study III: Access to setting and manager changes. We have hypothesized that more and frequent changes in settings and managers will promote more generalization. Therefore, following the record review and survey efforts, an observational study was undertaken to examine:

1. the frequency with which subjects encounter new managers (i.e., teachers, aides, therapists, etc.),
2. the frequency with which subjects accessed new settings,
3. the amount of time subjects spent in unsupervised activity,
4. the largest amount of time subjects spent with any one manager, and
5. the largest amount of time subjects spent in any single setting.

The data indicate that the percent of time subjects spent in a single setting ranged from 32% to 86%, with a median of approximately 66%. It may also be noted, however, that within settings, subjects spent relatively small percentages of their time with any one manager (about 30% on the average), and they generally were under some supervision or direction rather than simply being left alone. Average percentages of supervised time were greater than 70%, with a high of 96%. The frequency with which pupils encountered new managers and accessed new settings was considered impressive by the project staff.

The results suggest that pupils encountered a considerable variety of managers during the school day and that amounts of time spent with any one manager were within reasonable limits. The balance between amounts of supervised and unsupervised time generally appeared acceptable; however, one subject did spend more than 50% of the observation period in unsupervised, undirected activity. The conditions seemed favorable, on the whole, to the development of generalized responding across persons. On the other hand, it was found that, although relatively large numbers of settings were accessed, large portions of subject time were spent in a single setting (approximately two-thirds of the time, on the average).

Study IV: Peer interaction. Interaction between severely handicapped students and their peers is another ecological factor which may have an effect on skill generalization. When interaction does not occur, generalization (particularly of social and communication skills) might suffer. This study had two objectives. The first objective was to obtain an estimate of the amount of time during the school day in which subjects had the opportunity to interact with nonhandicapped or lesser handicapped peers. The second objective, derived from observations that spontaneous interactions between severely handicapped and lesser or nonhandicapped peers are infrequent in the absence of educational programming, was to determine whether managers prompt or reinforce interactions when opportunities for interaction exist.

The data indicate that the amount of time available for interactions in or out of school was highly variable and low, with medians ranging from 0% to 21%. Only 4 of the 10 subjects participated in community activities which provided opportunities for interaction, and the time spent in such activities was small (0.25 to 4.0 hours per week). Types of in-school activities with interaction opportunities included music, physical education, swimming, recess, and assemblies, while opportunities for interaction in the community included swimming and horseback riding.

Despite the dramatically low levels of interaction observed, it was found that managers typically failed to provide either cues or encouragement for interaction. Subjects were thus unlikely to interact with lesser handicapped peers and, due to a lack of educational programming, were considered unlikely to display increased involvement in the future.

Study V: Cross-setting generalization. The findings of the first four studies in this series indicated that elements of the educational ecology of subjects did not greatly favor the development of skill generalization across situa-

tions and persons. This study was conducted during the final weeks of, and immediately following, the academic year to determine through direct observation whether skills being taught to a subset of subjects did generalize to other settings.

Several of the findings were unexpected. The most surprising, however, was the frequency with which substantial degrees of cross-situational performance (i.e., generalization) were observed. Generalized responding was noted in 8 of the 10 programs studied. For four of the programs, correct responding occurred on 100% of the opportunities. It is possible that instruction within the school context is being undertaken on a variety of skills which pupils not only already have in their repertoires, but which are generalizing to other situations. This is underscored by the fact that impressive levels of correct responding were noted for three of the five programs which teachers thought were not generalizing. Subjects scored 100% correct on two of those programs and 80% correct on the third program.

Strong and significant relationships were found between average generalization and two manager and setting variables. The large positive correlation between new managers per hour and average generalization suggests that generalization may increase as pupils encounter greater numbers of managers throughout the day. The substantial negative correlation between generalization and the largest percent of time spent in any one setting indicates that subjects who were confined to a single educational setting (i.e., a single classroom) displayed lower levels of generalized responding than other subjects.

Discussion. Certainly, the findings of these studies must be considered both tentative and limited due to the small number of schools, subjects, teachers, and observations involved, the structure of observations, and the inaccuracy which might result from information gained from surveys rather than from direct observation. In many cases, however, the findings were so consistent across subjects, years, and sites as to instill a fair degree of confidence in the reliability and validity of findings in at least the participating sites. In the coming years, data collection will be extended to include not only a subset of the subjects who participated in the present series of studies, but additional subjects, teachers, and sites through a state-wide IEP survey. As data accumulates, extant conditions within the educational ecology of severely handicapped pupils should become more clearly defined, and critical variables more discernable.

Given caveats as indicated above, the data suggest some specific, although tentative, implications for structuring the educational environment of severely handicapped pupils in a manner which should increase the probability of generalized skill performance:

1. Specify generalization intent in IEP objectives.
2. Examine objectives to insure functional value.
3. Provide examples of relevant and irrelevant stimulus dimensions in training.
4. Communicate frequently with parents regarding pupil performance across settings and environmental requirements for skills being taught.

5. Expand opportunities for interaction with lesser and nonhandicapped peers in school and community situations and, when such opportunities exist, provide systematic encouragement.
6. Finally, provide manager and setting variation with each school day. Be particularly attentive to schedules for children displaying the most severe intellectual deficits as they may be at greatest risk for confinement to single settings for extended periods of time.

Summary of Second Year's Findings

The goals of the investigation and the nature of each study remained largely unchanged during the second project year. Five new subjects were added and the longitudinal nature of the study was maintained by continuing to follow five of the pupils in the original subject set. As in the first year, ecological conditions such as IEP content, manager and setting variation, interaction opportunities, and across-setting generalization were examined through record reviews, surveys, and direct observation. Additionally, in order to examine the generality of findings regarding the content of IEP objectives, the investigation was expanded to include an analysis of objectives from a variety of programs serving severely handicapped pupils from across the state of Washington. For more information on these studies, see Billingsley, Thompson, Matlock, & Work (1984).

Study I: IEP objectives and generalization. This study continued an analysis of IEP short term objectives to determine whether teachers indicated the intent to promote generalization across situations or persons, and whether the objectives were of such a nature that skill performance was likely to have adaptive value.

A total of 110 objectives were coded across the two sites. Overall, 3% of the total number of objectives coded for 1982-83 specified generalization intent, while a figure of 4% was obtained for 1983-84.

In terms of target behaviors coded as functional, obtained percentages were somewhat less than for either the 1981-82 or the 1982-83 school year. Whereas approximately two-thirds of the total number of behaviors in objectives were considered functional during previous years, only slightly greater than one-half were coded as functional during the 1983-84 year.

The findings of this investigation indicate that approximately one-half of the 110 IEP objectives targeted functional behaviors for instruction. Skills were being taught which might at least potentially contribute to adaptation and come under the control of natural maintaining contingencies. On the other hand, very few objectives, either singly or in combination, indicated generalized performance as a desired instructional outcome. That finding is consistent with data reported during the previous project year.

Although the present findings do not provide evidence regarding the degree to which teachers who were involved in IEP development actually programmed for or evaluated generalized performance, they do indicate that objectives did not include information which would guide program man-

agers toward effective programming and evaluation practices. The findings also suggest that approximately one-half of the skills identified for instruction in IEPs may be of questionable adaptive value and, as a result, might fail to generalize or maintain outside of the training setting. A failure to obtain generalization could be due to both a negligible demand for such behaviors in nontraining environments and to the low density of natural reinforcement which may follow the performance of nonfunctional skills.

Study II: An examination of the contents of a statewide sample of IEPs. The findings of Study I must be interpreted cautiously due to limitations related to sample size and type. It is possible that different results would have been obtained had a broader sample of IEPs prepared by more teachers been examined. Results might also have been different if IEPs for pupils from integrated rather than segregated facilities had been included in the sample. The purpose of this study was to determine the extent to which objectives included in a statewide sample of IEPs from programs that served pupils with severe handicaps specified generalized outcomes and targeted functional skills for instruction.

Each year, the Office of the State Superintendent of Public Instruction (OSPI) for the state of Washington reviews a sample of IEPs from districts across the state. During the 1983-84 school year, OSPI reviewed IEPs within 78 districts, the smallest of which served 40 students and the largest of which served over 26,000. It was from among those districts that 60 IEPs from 14 different districts were obtained for coding by project staff. Four of the school districts contributed IEPs from segregated facilities; all districts but one provided IEPs from integrated facilities.

A total of 618 objectives were coded across all districts. The data from the statewide survey are highly consistent with data presented in Study I in terms of the high degree of variability which existed in the ranges of total numbers of objectives included in IEPs and overall percentages of objectives considered as functional. Median number of objectives included in IEPs were also highly similar to those presented in Study I. It was surprising, and rather disheartening to note, however, that the lower end of the range of total objectives included in integrated setting IEPs was zero. In other words, after objectives which specified deceleration targets, required only passive participation, or specified training only for evaluation activities were excluded from the coding process, one student was left with no skills explicitly targeted for instruction.

Overall, the percentage of objectives within the statewide sample which specified a generalized outcome was greater than that obtained in Study I. However, the percentage was small and the IEPs for some pupils contained no objectives indicating the need for generalization; in fact, although the overall percentages of objectives specifying a generalized outcome were quite similar for both integrated and segregated settings, the average number per IEP within integrated settings was zero.

Percentages and median numbers of objectives which targeted functional skills were similar in both integrated and segregated settings (i.e., approximately 50% with medians of 3 and 2 objectives respectively). Although the

data suggest that some IEPs prepared for pupils in integrated settings contain more functional objectives than the highest number of functional objectives contained in any IEP prepared for a pupil in a segregated setting, both settings contributed IEPs in which no objective was coded as functional. Seven IEPs from integrated settings (14%) contained no objectives which were considered functional and one IEP from a segregated setting (9%) contained no functional objectives.

This study tends to confirm the conclusions presented in Study I. That is, IEPs collected from across the state of Washington included few objectives which specified generalization as a desired educational outcome. Further, although the number of objectives related to functional skills was considerably larger than the number related to generalized outcomes, a substantial percentage of objectives (in this case, greater than 50%) targeted skills which could not be considered functional according to the criteria employed. The sites which contributed IEPs for Study I, therefore, do not appear to be particularly deviant in terms of degree of functionality reflected within instructional objectives. However, the percentage of objectives which specify generalization may tend to be somewhat smaller at those sites than is generally the case within the state. An additional finding of this study was that integrated and segregated facilities did not appear to differ dramatically in average number of objectives included in IEPs, or in average number or total percentage of objectives coded as specifying generalization or indicating functional skills.

Study III: Parent survey. Parents/guardians of the 10 subjects in local districts were surveyed to determine whether they thought that objectives written for their children were functional, whether pupils had the opportunity to perform skills in nonschool settings, whether skills were performed at appropriate times, whether skills were performed at inappropriate times, and whether home training was provided.

Parents were highly consistent in perceiving objectives as clear. In other words, they indicated an understanding of what the schools intended to teach. They also indicated, as did parents during the previous project year, that they considered the vast majority of the behaviors taught to be functional.

In contrast, parents reported that their children had the opportunity to perform a considerably smaller percentage of objectives outside the school. Whereas 94% of the objectives were considered functional, opportunity to perform was noted in the case of only 57%.

Parents at each site reported that their children performed 45% of the skills specified by the objectives at appropriate times outside of the school, suggesting the presence of at least some degree of across-setting and across-person responding. For both sites, however, the percentage was less than that reported for the previous year.

A relatively small percentage of skills was reported as occurring at inappropriate times, and home training was reported for approximately the same percentage of skills as those which were indicated as being performed

appropriately. The latter finding may, of course, suggest that home training plays a significant role in the establishment of generalized skill performance. In terms of comparisons across years, home training was reported to occur at exactly the same level at Site 1 for the 1982-83 and 1983-84 project years, but a decrease of 17-percentage points was observed at Site 2. For continuing subjects, the percentage of skills on which home training was reported decreased at both sites.

Study IV: Access to managers and settings. This study was implemented to determine whether subjects accessed multiple settings and encountered multiple managers during the school day or whether they were confined largely to single classrooms and contacted by one or very few managers (e.g., teachers, aides, and other service providers).

The second year data indicate that the percent of time subjects spent in a single setting ranged from 25% to 77%, with a median of approximately 46%. It may also be noted that, within settings, subjects spent relatively small percentages of their time with any one manager, with a median of less than 24%. Consistent with data obtained during the previous project year, subjects encountered new managers and accessed new settings with impressive frequency.

Although subjects seemed to encounter a considerable number of managers, it may be noted that, on the average, considerable periods of time were scored as being unsupervised. While one subject at each site was supervised for over 90% of the observation period, the median amount of time supervised was less than that reported for the previous project year.

In summary, most of the manager and setting variables examined in this study seemed conducive to the development of generalized responding across settings and persons. An obvious exception, however, was observed in average percent of time supervised. It seems probable that, where considerable amounts of time are spent in unsupervised conditions, the opportunity to practice and receive reinforcement for newly acquired skills will be limited. The result of such practice and reinforcement limitation could be a failure to maintain and generalize skills. Of course, if pupils were unsupervised, but interacting with other pupils and performing skills which would elicit reinforcement within the interaction context, the generalization and maintenance of some skills might be enhanced.

Study V: Peer interaction. Based on the first year studies, it was suggested that, where few opportunities for interaction exist, generalization of social and communication skills might suffer due to (a) insufficient exposure to exemplars of persons or situations which would control appropriate responding, (b) the inability to encounter natural maintaining contingencies, or (c) limited occasions for managers to engage in sequential modification. The objectives of this study were the same as those indicated for Study IV, "Peer Interaction," of the first project year.

As was the case during the previous project year, data regarding time available for interactions was highly variable across both days and subjects. Median percentages of time available within school ranged from 0% to 96%

Washington
Research
Organization

ERRATA

Investigating the Problem of Skill Generalization
(3rd Edition)

Page 33. Sentence beginning on line 3 should read, "In terms of comparisons across years, home training at Site 1 dropped 8-percentage points from the 1982-83 to the 1983-84 project year and a decrease of 3-percentage points was observed at Site 2."

Page 34. Sentence beginning on line 1 should read, "At each site, one subject was reported to have no available times for interaction, and the median amount of time available for 4 subjects was reported to be zero."

PLEASE NOTE: Washington Research Organization (UWRO) has moved to new offices. Our telephone number remains 206/543-8565. Our new address, effective immediately, is:

Washington Research Organization (UWRO)
110 Miller Hall; DQ-05
University of Washington
Seattle, WA 98195

across all subjects. At each site, one subject was reported to have no available times for interaction, and the median amount of time available for 5 subjects was reported to be zero.

Data regarding the median total number of interaction intervals scored at each site indicate that, on the average, very little interaction occurred between the subjects and their peers. Although the number of subjects involved in interactive activities was greater in 1983-84 than noted during the preceding project year (8 as compared to 4), the average length of time per session that subjects could have been engaged in interaction was quite short. On the other hand, the range of total intervals scored, both across subjects and across observational sessions, was broad at both sites. No interaction was observed in the case of some subjects, while other subjects were involved in interactive activity for more than half of the 75 intervals during at least one of the two observational sessions. It should be noted, however, that where substantial levels of interaction were recorded, such levels were not necessarily maintained consistently across sessions.

Data indicated that, based on teacher reports, some subjects spent very small proportions of their time in situations that would provide the opportunity for interactions with less- or nonhandicapped peers. Other subjects spent relatively large blocks of time in such situations on some days. Substantial and consistent opportunities were most apparent where service delivery was provided within a combined classroom structure which provided the opportunity for pupils of varying degrees of disability to be "mixed" for a relatively large proportion of the school day. Teachers reported that very little time was spent by any subject in community-based activities. During the observations of interactions which occurred when opportunities were provided, the most general outcome was few interactions of brief duration. Where substantial levels of interaction were noted, they were likely to be sporadic. Finally, managers were never observed to encourage peer interaction.

Although it may be that teacher-directed structure, under at least some conditions, may be ineffective in promoting or even hinder interactive behaviors, the low levels and sporadic nature of interactions noted in this study suggest that instructional interventions might be highly appropriate. No revision of the first year conclusions seems to be warranted by the data obtained in this study.

Study VI: Skill generalization. It was reported during the first project year that, although elements of the educational ecology did not seem to favor the development of generalized skills, considerable cross-setting performance was noted. The conditions under which generalization assessments were conducted, however, were highly structured. That is, parents were taught critical features of program administration in terms of the nature of directions and the appropriate level of antecedent assistance to be provided for each skill step. If the teacher used a gesture to prompt a pupil to pull his socks up above his ankles, the parent would be taught to employ that level of assistance for that particular step during generalization probes. Responses were then considered "generalized" if they were emitted when the designated level of assistance was applied by the parent.

The present investigation was designed to (a) reduce the probability that skills selected for observation were already in the repertoire of the subjects and (b) allow us to determine whether generalized performance would be observed when pupils achieved the maximum level of independence desired by the teacher (i.e., the "mastery level") for a particular skill.

It may be noted that each of the 5 subjects who was reported to have achieved mastery of skills within the school setting did, in fact, perform at mastery level when final assessments were observed by the project staff. It may also be noted that 4 of the 5 subjects who achieved mastery in the school setting also displayed mastery levels of responding in the generalization settings, with 2 subjects displaying slightly more proficient skill performance within the generalization setting.

The 4 subjects who displayed generalized performance were a diverse group in many respects. They ranged in age from 10.75 to 17.25 years. Three subjects were male and 1 was female. One was in a primary program, another was in an intermediate program, 2 were in secondary programs, and all had different teachers. They spent between 56.2% and 93% of their time in supervised activity, and between 13% and 48.3% of their time under the supervision of one manager. Between 30% and 61% of their time was spent in one setting. In terms of total number of interaction intervals scored, however, all had relatively few (or no) interactions. Near the end of the study, parents were asked whether their children had received home training on the target skills. Three parents replied that such training had been provided and one indicated that it had not. Given such diversity, subject or ecological variables which might account for the observed generalized responding are not apparent. There remain, however, at least two additional factors which might relate to generalization: sites and skill functionality.

With regard to sites, although 4 of the 5 subjects who mastered their objectives were from Site 1, one of those Site 1 subjects did not display any degree of generalized responding. That fact, combined with the small data set available for analysis, does not permit one to conclude that sites, per se, were related to the frequency with which generalized performance was observed. With regard to skill functionality, the single subject who displayed a 0% level of independent performance in the generalization setting was the subject who had been instructed in the only skill which had not been coded as functional by the project staff. This finding is at least consistent with the correlational data provided in Study III which tends to support the contention that skill functionality may promote cross-situational performance.

In conclusion, where teacher-specified mastery levels of performance were achieved, cross-situational mastery level responding was observed in all but one instance. This rather impressive frequency of generalized performance occurred in spite of the fact that none of the teachers had specified generalization as a desired educational outcome in their instructional objectives.

General conclusions. By and large, the findings from the second year studies leave one with the same impression as those reported during the first project year: The educational ecology of the subjects was not generally structured in a manner consistent with the principles of generalization

outlined by Stokes and Baer (1977). IEP objectives did not appear to be written with generalization in mind (at either the local or state level), parents supported the provision of instruction related to skills of doubtful functionality, subjects often spent large blocks of time unsupervised, few opportunities were made available for interaction between subjects and members of the community at large, and encouragement for interaction with less-handicapped and nonhandicapped peers was never observed.

Although some of the findings appear to paint a rather bleak picture, all but 1 of 5 subjects who mastered target skills during the school year were observed to perform those skills within generalization situations and parents reported cross-situational performance in the case of a substantial percentage of the skills identified in IEP objectives. Many factors which were not the subject of this investigation (e.g., instructional methodology) may have contributed to the observed or reported levels of generalized responding. However, the limited data collected to date seem to suggest that skill functionality may be a variable which is at least deserving of continued, and closer, examination. During the coming project year, increased attention will be given to that variable, particularly within the context of actual observation of cross-situational responding.

Intervention study: Effects of competing behaviors on generalization. The existence of old behaviors within an individual's repertoire may compete, and thereby interfere, with the generalization or maintenance of newly acquired, more desirable behaviors. Response competition might occur because a desirable behavior fails to produce reinforcement as reliably as the less desirable behavior, because the less desirable behavior permits more efficient access to reinforcers, or because of the lengthy reinforcement history associated with undesirable behaviors (e.g., tantrums) and the resultant strong stimulus control exerted by stimuli in nontraining settings over those behaviors.

The potential problem which competing behaviors may pose for generalization was noted by Stokes and Baer (1977) in their discussion of the "introduce to natural maintaining contingencies" (p. 350) principle of generalization programming. In that discussion, it was suggested that rearrangement of nontraining environments may be necessary to insure that undesirable response forms are unsuccessful in gaining access to those reinforcers which may also be obtained by newly learned, more desirable forms taught in a training setting.

This study had two purposes. The first was to examine the extent to which competing behaviors interfere with the generalization and maintenance of new behaviors. The second purpose was to investigate the effectiveness of reinforcement denial (i.e., the rearrangement of maintaining contingencies) as a tactic to facilitate generalized responding in situations in which both old and new responses serve the same function.

Two pupils (Paul and Helen) with severe behavior disorders and mental retardation served as subjects for the study. The children were enrolled in different classrooms for pupils with behavior disorders at a university affiliated campus school. They were nominated by their classroom teachers for

inclusion in the study because each displayed a long-standing behavior which was undesirable to the teacher, which seemed to serve a relatively obvious function, and which the teacher wanted to replace with a more appropriate behavior. For both of the subjects, the general approach involved assessing the subject's use of the new and of the old behavior throughout a *series of phases within an "instructional"* (i.e., classroom) setting and within one or more generalization settings in the presence of a manager other than the instructional setting manager. Assessment within the instructional setting occurred during the regularly scheduled lunch period. Generalization assessments were conducted at snack times.

It was found that behaviors which served a relatively obvious function, and which were desired by classroom teachers, were replaced by less desirable behaviors which served the same function within maintenance and/or generalization situations. It was further demonstrated that the desired behavior could be elicited in those situations by reducing the functionality of the less desirable response. Both general and specific implications for educational practice are discussed.

The results have both general and specific implications for the promotion of skill generalization across settings and managers. Generally, it appears that where new behavioral forms are being taught which will serve a new function, generalization may be hindered by a failure to employ procedures designed to increase similarities across (or decrease discriminability between) training and generalization situations. When new forms are being taught to replace old forms which serve the same function, the above may also be true; however, the old form may also interfere with the appearance of the new form. In terms of specific implications for teaching, where old forms compete with new forms, it may be unnecessary to reinstate training procedures in new situations. Rather, it may be sufficient to reduce the functionality of the old forms by ensuring that they do not permit access to reinforcement. (Alternately, although unexplored in this investigation, generalization might be enhanced by increasing the efficiency of the new forms.) In addition, failures of generalization may often reflect a failure, or inability, to deny access to reinforcers in nontraining situations, rather than a programmatic deficiency in training settings.

Intervention study: The effects of trial sequencing on generalization. Recent studies show that handicapped children can acquire new skills when taught within a total-task, distributed trial framework, but there is no evidence that this strategy facilitates generalization. Instruction presented in a massed, multiple trials format, a traditional instructional strategy, has also been successful in teaching severely handicapped children a variety of skills, but there is conflicting evidence regarding the effects of such a strategy on skill generalization. The present study was designed to evaluate the effects of traditional and functional models on the acquisition and generalization of a functional task. Specifically, instruction presented within a total task, single trial instructional format was compared with a backward chaining, multiple trials procedure in terms of the acquisition and generalization of an 11-step task involving snack preparation by children with severe behavior disorders.

Five male and three female children, ages 6.9 to 11.4 years, served as subjects. The instructional strategies used in this study included several specific instructional techniques. Multiple trials instruction, comprised of procedures commonly utilized with severely handicapped learners, included the presentation of multiple trials, serial task sequencing presented in a backward chaining format, and "artificial" training conditions. Instruction within a total task framework consisted of distributed trials, presented within a total-task (i.e., concurrent chain) format, and in-context instruction. Under the total task format, subjects received only one trial daily.

Results indicated that 6 of 8 subjects acquired more independent steps under total task training than under multiple trials training. The remaining 2 subjects failed to acquire any independent steps under either experimental condition. Superior net gains in levels of assistance (i.e., changes toward greater independence) were made under the total task condition for 6 of 8 subjects. One of the remaining subjects made a net gain of one level of assistance under both experimental conditions; the other subject made a slightly greater net gain with multiple trials training than with total task training. The acquisition data also show that the total instructional time for total task training was only 29% of the total instructional time for multiple trials training.

Preliminary analyses of the data indicate: (a) no appreciable difference between the generalization scores following total task and multiple trials instruction for 3 of 8 subjects; (b) one subject's score on one generalization probe was not indicative of his usual performance; disregarding his third probe, his data reflect no difference between the degree of generalization following total task and multiple trials training; (c) one subject evidenced superior generalization following the total task condition; and (d) for 3 of 8 subjects, probe data following both conditions were equivocal.

Studies in Performance Patterns

Most people agree that each pupil is an individual and that what might work with one student may not work with another. There is a need to individualize. Individualization usually begins with an identification of the skill areas and behaviors to be taught. Next, detailed inventories of the pupil's skill in each area of the curriculum are conducted to determine, for each behavior selected, the exact level or curricular step at which instruction should begin. Major pupil characteristics which might indicate the need for a particular instructional approach are also identified. The teaching procedures which might prove most effective with mentally retarded children, for example, might be quite different from those which work best with the deaf-blind. Surveys of "learning channels" and "reinforcement preferences" could also be used to help in the development of specific instructional plans. Overall, there is much that can be done to select and develop highly individualized approaches for meeting a pupil's needs.

For the most part, however, educators tend to think of the factors which might determine the effectiveness of an instructional approach as being rather fixed and unchanging. The student is and always will be deaf-blind;

the student is "visually oriented," or "prefers juice instead of hugs." In reality, instructional approaches which work quite well on one day may hinder further learning on the next. Truly individualized instruction will involve the continuous assessment of daily pupil progress to determine exactly when and how instructional procedures should be modified to keep pace with the changing needs of the pupil.

Fortunately, research over the past decade has identified patterns in the way pupils' learning changes from day to day. Each pattern can be related to specific instructional needs. For example, there is a surprisingly consistent relationship between a pupil's overall fluency in performing a task and the need for additional guidance. If a pupil is performing a task very slowly (even if overall accuracy is fairly good), strategies such as increased cues, prompts, and corrective feedback may facilitate further progress. However, if the pupil is performing the task fairly quickly, those same strategies may be quite ineffective. After noticing the ineffectiveness of one strategy, teachers may need to try three or four different approaches before finding one that works. Of course, soon after finding one that works, the pupil's needs change once more and the process of trying to find effective instruction begins all over again.

By examining the performance characteristics of students who were acquiring or building fluency in a skill, researchers found certain elements of performance to be very important. These included the student's correct rate or frequency of performance, the accuracy of the response, the weekly rate of learning or progress, and the variability of responding. Five constellations of these elements were identified as specific performance patterns. By examining those characteristics, researchers were able to predict whether or not a specific strategy would help or hinder the student's learning. To replace the guesswork in programs for acquisition and fluency-building, rules were developed to help teachers match instructional strategies to changing pupil needs. Research shows that teachers who follow the rules are able to choose an effective strategy ten times out of twelve.

As successful as the performance pattern rule research has been, to date it has only looked closely at the way in which pupils learn and master new skills in specific instructional situations. Very little is known about the relationship between those patterns of learning and the chances that the new skill will generalize to other situations. It will be the purpose of the performance pattern research at UWRO to investigate those same elements to discover their relationship to generalization. We will then try to match specific patterns with instances and noninstances of generalization. If necessary, we may look at other elements in responding, but we hope that the same elements will prove predictive of generalization. We will attempt to identify the instructional procedures with the highest probability of promoting generalization. If we are successful, we will be able to match particular types of instruction with student's individual needs in order to facilitate generalization.

Design of Descriptive Studies

A great deal of potentially useful information concerning the relationship between patterns of learning and generalization already exists. For example,

a published research study, originally conducted to determine the usefulness of feedback in promoting generalization, might be evaluated to look at the relationship between performance patterns and generalization. Similarly, the data already being collected in many classrooms to monitor pupil progress may yield certain clues. There are at least two advantages in using existing data—it is far less expensive and there is no chance that our expectancies of what should happen might somehow affect what does happen. The disadvantages in using existing data lie in the fact that they may not provide all of the information required for the study (e.g., most researchers have expressed performances in terms of simple percentages, or accuracy statements, rather than in both accuracy and fluency as desired for the current research), and some questions often exist concerning the reliability of the data and/or the exact nature of the procedures employed to facilitate generalization. To date, 115 research studies have been evaluated, and have provided at least some information of use to the Performance Patterns research. To supplement those data, however, it became necessary to begin direct observations of our own during the second project year.

Experience in special education research has shown that valuable data can be obtained from scientific observation of what is already happening in the classroom before making any changes. During this phase of the research, the project has been monitoring and documenting what is already going on and how those activities appear to relate to generalization. Severely handicapped pupils with a wide range of disabilities have been included in the study. In previous performance pattern research on skill acquisition and fluency, basic pupil characteristics (i.e., type and level of handicap, age, sex, etc.) were not strongly related to the way in which performance patterns predicted the success of various instructional approaches. Nevertheless, detailed records of pupil characteristics are being kept and evaluated to determine whether those characteristics do relate to the usefulness of performance pattern rules in predicting when and how generalization might be facilitated.

Teachers volunteering for the study collect (and allow project staff to collect) specific information concerning daily pupil progress in a sampling of instructional programs. Concurrently, project staff monitor each pupil in a variety of other situations to determine if, when, and how the pupil begins to demonstrate new skills outside of the instructional setting. General observational studies have provided the research staff with the information needed to refine specific hypotheses, and have provided the basis for more directive interventional studies.

Design of Intervention Studies

Beginning during the latter part of the second project year, specific studies are being conducted to clarify the relationship between performance patterns in an instructional situation and the likelihood of generalization. For example, noninterventional studies conducted earlier suggested that pupils who achieve a specific level of fluency in the instructional setting are more likely to generalize their skills, so a study is now being conducted to test that relationship by bringing skills up to a level of fluency and noting whether

generalization actually does occur. When a reasonably comprehensive set of rules has been developed, the impact of those rules will be tested by training new teachers in their use and evaluating the effect of rule use on generalization. In a second study, the effects of different types of stimuli and reinforcers to eventual generalization are being studied. Results to date suggest that "natural" consequences can indeed be effective in developing new behaviors, but are not entirely effective in facilitating generalization in the absence of natural stimuli as well. Direct intervention studies will continue throughout the third and fourth project years.

During earlier performance pattern studies, the success of a program change was judged by the immediate impact on performance, the change produced in average weekly progress, and the net effect of those two factors on eventual skill mastery. Those same variables are being used to monitor the basic effectiveness of any changes made in the instructional situation to improve generalization, but special probes of the pupil's behavior in a variety of other situations are also conducted to examine generalization. Initially, the degree of generalization at any point in time is being described in terms of the number and type of noninstructional situations in which the behavior is observed to occur, and the degree to which performance characteristics in the noninstructional setting approximate those observed in the instructional situation (in terms of fluency, accuracy, and improvement over time).

Expected Outcomes and Products

If the proposed studies are as successful as earlier work, it should be possible to develop a set of rules which teachers can use to evaluate individual pupil performance and decide if, when, and how they might change instructional procedures to facilitate generalization. Rather than impose a single approach to developing generalization, the rules would help teachers to choose the best method, from among a variety of possible instructional procedures, to meet the individual needs of a pupil at a given point in time. With such rules, it will be possible to truly individualize instruction to take into account each pupil's changing needs.

In addition to a series of research papers and monographs documenting the progress of individual studies, the performance pattern research should result in the creation of a brief "user's manual" which explains how the rules can be used to facilitate skill generalization with severely handicapped pupils. The manual will be written in a manner which is easily understandable to teachers and other educational practitioners and will be as self-contained as possible. The manual will not assume that the reader has any prior knowledge of the skills necessary to use the rules. The actual usefulness of the manual will be tested on a group of teachers toward the end of the fourth project year. The feedback gained from that trial implementation will be used to make modifications during the fifth and final project year.

Summary of Studies to Date

Our first step was to review the existing literature to identify strategies that had a high probability of promoting skill generalization. In addition, studies

that presented precise individual and repeated measures of skill generalization were identified to begin the process of identifying parameters of performance that might be linked to the matching process.

The general procedures for the retrospective analysis of published data have included the development of a form for coding information about the studies, establishing coder reliability on the content and format of the coding form, reading and coding articles, entry of coded information in a computer, summarization of the data collected, and analysis. During the first and second years, one to three research assistants worked with Dr. Norris Haring and Dr. Owen White on the Performance Pattern Studies. For more information on these studies, see White, Haring, and Miller (1983) and Chapter 5 of this publication (i.e., White, Leber, and Phifer, 1985).

A number of basic hypotheses had been formulated during FY 82-83 concerning possible relationships between patterns of performance during training and eventual skill generalization (or failure to generalize). Rather than engage in a series of focused experimental studies during FY 83-84, it was decided to test the basic validity of those hypotheses through extensive monitoring of whatever functional behaviors were already being taught in participating classrooms. Unfortunately, observation of ongoing instructional programs encountered a number of difficulties which severely limited its usefulness to overall project goals.

The primary difficulty in observing extra-instructional behavior lay in the simple fact that most of the observational targets could not be emitted without access to special equipment, materials, or people—access which was not always arranged and/or was arranged only infrequently or for brief intervals. For example, the behavioral target in a drawing (prewriting) program for one subject required access to paper and crayons; another subject's precommunication switch-activation program required access to a tread-switch and tape recorder; the behavioral targets for a self-feeding program could, of course, only be assessed during meal and snack times; and natural opportunities for another program, "hands up" to allow the wheelchair tray table to be put on/removed, were simply not very frequent. Classroom staff were very cooperative in attempting to increase the opportunities for extra-instructional behavior, but several factors (e.g., equipment breakdowns, extended subject illnesses, subjects leaving early for vacation) continued to make meaningful observation difficult. Overall, it was simply not possible to collect sufficient amounts of reliable data through general observation to test study hypotheses.

To correct those difficulties, pilot studies were designed and implemented to test specific hypotheses. These experimental studies will continue through the next year, concentrating on the role of fluency in determining the probability of skill generalization, especially in cases where competing behavior patterns already exist within the subject's repertoire.

A second series of studies has attempted to investigate the impact of "natural consequences" on skill acquisition and generalization. In theory, if the consequence for an act is directly related to the act (e.g., getting to play with an item named, rather than being reinforced with an edible), the probability

of skill acquisition and generalization should improve. Earlier studies have not yielded clear results, however, due to a number of methodological flaws. In the study conducted by the Patterns Research Team, we have attempted to correct those difficulties and to develop new procedures for monitoring the acquisition/generalization process in a way that will allow much more to be operating at the time a skill first emerges and the time when it first begins to generalize.

Preliminary results from those studies suggest that: (1) natural consequences (e.g., being allowed to play with an item identified in a receptive language program) are effective in developing language discrimination skills, but do not produce rates of acquisition quite as rapid as the use of "artificial" consequences (e.g., edibles for pointing to a named object); (2) once acquired, skills generalized well following either reinforcement condition within stimulus dimensions (e.g., across different photographs of objects), but not across stimulus dimensions (e.g., from photos to objects) regardless of the consequence used; and (3) when training stimuli were varied (e.g., objects were used for stimuli, rather than photographs), generalization across stimulus dimensions (e.g., photos and objects) improved when artificial consequences were used, but not when natural consequences (i.e., opportunity to use the object) were used. Overall, then, consequence type did seem to effect both rate of acquisition and probability of generalization, but not in ways that were initially hypothesized. A more detailed evaluation of the data may help in identifying possible controlling factors more clearly.

Pilot Study: Investigating Strategies for the Systematic Fading of Cues, Reinforcers, and Contingencies

Artificial and exaggerated prompts and cues have been established as powerful components of instructional procedures in the education of severely handicapped students, but these procedures may inhibit successful generalization of the instructed behavior across settings. Cue-fading, the gradual elimination of these artificial prompts to allow the behavior to come under the control of natural cues (and thus improve generalization), is recommended to teachers of the severely handicapped, but rules for determining precisely how to fade cues have not been discovered. Thus many teachers wait until students have reached 100% mastery to begin the fading process. This may create an overdependence on non-naturally occurring cues and may actually inhibit generalization. The purpose of this specific pilot study was to investigate one set of systematic strategies for cue-fading. The study was conducted by Greg Weisenstein, Sharon Field, and Carol Kiolet. For more information on this study, see Weisenstein, Field, & Kiolet (1984).

This pilot study addressed the following research questions: (1) Will fading in the training setting increase or decrease the frequency of target behaviors in the natural setting? (2) Can fading be initiated at or prior to 80% mastery in the didactic setting without a negative side effect on generalization of skills in the natural setting?

The subject was 18.5 years old. He attended a secondary special education program for handicapped youth, located in a vocational-technical institute (VTI). School records show administration of the Wechsler Adult Intelligence Scale (Wechsler, 1955) seven months prior to the start of the study resulted in a verbal IQ of 58, performance IQ of 54 and a full scale IQ of 54. Primary problems in social behavior included off-task and failure to initiate contact with authority figures.

It appears that correct responses in the classroom setting during the *Classroom Training With Prompts* condition were increasing, but stabilized at the introduction of the fading procedures. The stabilization was at the correct response rate of the final observation in the previous condition. While apparently the upward trend of correct responses was interrupted, there was no decrement observed during the *Initiate Fading* condition. This indicates that if level of performance in the training environment is a concern, fading procedures should not be instituted until an acceptable level of performance in the training environment has been established.

Natural environment data reveal an upward trend in *Baseline*, followed by a drop at the introduction of the *Initiate Fading* condition. The upward trend was again repeated during this condition, with a drop at the introduction of the *No Prompts* condition. Again, the drop was followed by an increase in correct responses during this phase. During *Baseline*, there was an indication of increasing errors, but there were no errors during any subsequent phases. Therefore, the fading procedures were not observed to increase error responses, and may have even reduced error responses.

While the rate of correct responding initially fell during each phase, a strong recovery was observed as each phase progressed. As the initial decline in the rate of correct responding in the natural environment was consistently regained, and the performance level in the training setting was stable, it does appear that fading procedures may successfully be instituted prior to 80% mastery. If fading procedures can be instituted prior to 80% mastery, it may provide for greater generalization as the student does not become as dependent on non-naturally occurring cues. It also may help to promote faster and more efficient learning.

Conclusions and recommendations. The results of this pilot study provide insight regarding the directions of future investigations of fading techniques to promote social skills generalization in severely handicapped students. While the use of fading is independent of any particular behavior, it seems especially suited to the subtleties involved in social behavior. Because the naturally occurring cues and reinforcers for social behavior are often very subtle, it is often necessary to use exaggerated prompts and reinforcers to cue and reinforce the desired behavior. The results of this study support the hypothesis that generalization may be enhanced by fading the exaggerated cues to approximate the naturally occurring cues as closely and as quickly as possible.

The generalization demanded of the subject in this investigation was extensive, in order to remain as "true to life" as possible. The subject was asked to generalize across several settings, people, and cues. The indications of the

occurrence of generalization are therefore quite promising. Also promising is that, while not systematically investigated, the examiners were alert to the possible occurrence of increases in other undesirable behaviors as the subject's incorrect response rate in the target behavior decreased, yet noted that none occurred.

Future research should investigate the use of fading techniques for promoting generalization, with the ultimate goal of establishing a "formula" for fading. Investigations should look not only at when to begin fading, but also at the number and size of the steps which should be designed for the fading procedure. This pilot study indicates that to institute fading prior to 100% mastery may increase efficiency in the acquisition of social skills as well as facilitating the generalization of those skills across settings, people, and time. It is hypothesized that fading prior to 100% mastery may decrease the dependency on non-naturally occurring cues and thus promote generalization. The results of this pilot study indicate that this is an extremely important area for future research.

Studies in Self-Control

Typical instructional procedures for skill acquisition and fluency-building rely almost exclusively on a teacher or other trainer acting as the focal point. In almost every research and/or curriculum report, the handicapped person is seen as the one whose behavior is to be changed, rather than the individual who is to change her own behavior. This emphasis is evident when you consider that in most training programs:

(1) The behaviors to be changed are selected by others.
(2) The training materials and procedures are selected by others.
(3) The training procedures are implemented by others.
(4) Changes in behavior caused by training are monitored by others.
(5) Decisions about changes in training procedures are made by others.

Although this instruction has been effective in teaching specific skills, the collective effect of many years of such training may be to teach the handicapped individual total dependence on others for control in each situation. Generalized responding may fail to occur simply because the individual is waiting for someone to give step-by-step instructions in what to do.

Self-control procedures offer an alternative. In self-control training, individuals are taught how to use different techniques to direct their own behavior. It is easy to find examples of self-control techniques in everyday activities. One common self-management procedure is self-monitoring, or counting the occurrence of one's own behavior. A person who says, "This is only my third cigarette today," is monitoring her own behavior. We've probably all heard someone say, "I'm getting fat, I'll skip dessert." Such individuals are not only monitoring their own behavior; they are making a decision based on the information as well.

Another technique we use to manage ourselves is called self-instruction: directing the sequence of activities we are performing or are about to perform. People facing several different tasks or a particularly complicated task will often audibly list, to themselves, the sequence of things they are going to do. For example, "I'll start the water for the noodles, then I'll cut up the asparagus, then I'll put the noodles in, next start the asparagus, and hope that they are finished cooking at the same time." A third typical procedure is self-reinforcement, including selecting and delivering consequences for activities. For example, a person may reward himself with time to read the newspaper after he has washed the dishes.

While instances of these activities abound in our daily lives, until recently, little research in self-control has been reported. It is known that many people do not learn to use self-control skills without direct training in the skills. Research does show that self-control skills are usually just as effective as external-control procedures in changing behavior. Moreover, self-control may be better at facilitating maintenance and generalization, since the individual learns independence, rather than dependence.

Can we teach self-control skills to persons with severe handicaps? Only a few researchers have worked with handicapped individuals in teaching self-control skills, so this question has not yet been answered. We can develop empirical studies to determine if precise skills, such as pushing a button on a counter following task completion (i.e., self-monitoring), can be learned. We can also determine if other self-control activities help the person to change her own behavior and if they are effective in changing other behaviors in new settings. A second puzzle for research concerns the nature of the training. If the methods used to teach self-control skills rely on an external agent, will that method counteract the development of independent control? What other types of training can be used?

The purpose of research in this area is to investigate whether or not severely handicapped people can be taught to use methods of self-control. If so, what are the best methods of training? And if the self-control skills can be used by individuals to change their own behavior, do such skills improve generalization?

Design of Studies

The variables that will be investigated in these studies will include:

(1) the accuracy and fluency of the performance of the self-control skill,
(2) the length of time required for acquisition and fluency of the self-control skill,
(3) the instructional procedures used to teach self-control, and
(4) the effect of the self-control skill on the target behaviors (i.e., does the self-control skill facilitate generalization).

Although it is difficult to predict the course of future research, we will attempt to investigate each of the three primary self-control skills: self-monitoring, self-reinforcement, and self-instruction.

Three other self-control techniques will be integrated into the studies: self-determination of behaviors for change, self-determination of consequences, and self-determination of the ratio of behaviors to consequences. Subjects will be able to select consequences and behaviors to change in most instances, so that they can immediately begin participating in the behavior change process. Individuals who do not respond to questions (e.g., What would you like to work for? What will you do if you make a mistake? How many do you want to earn?) will be presented with a multiple choice situation via pictures, objects, or words during each training session.

Each study will include several different phases for the self-control behaviors as well as for the target behaviors (i.e., the ones we hope to change or affect by the self-control skill). Repeated measurement data will be collected on the self-control and target behaviors in training and nontraining settings. Data will be summarized and analyzed according to the accuracy and fluency of the response, and by changes in the individual's level and direction or trend of performance. Data will be collected in training and in nontraining settings, with the subjects' regular teachers and with persons unfamiliar to the subject.

Following the collection of baseline data, the subject will be taught to use a self-control skill by her teachers. The effect of the self-control skill on the target behavior will be measured (a) to see if and how the performance of the target behavior in the training setting and with the teachers is affected, (b) to see if and how the performance of the target behavior outside of the training setting and/or with nontrainers is affected, and (c) to see if the self-control skill is generalized across settings and/or across behaviors. Opportunity to overtly apply the self-control behavior may then be withdrawn (e.g., by taking the counter away from the subject).

Expected Outcomes and Products

Since there are so few precedents for teaching self-control to severely handicapped students, it is difficult to predict the sequence and nature of the studies. Findings from one study will probably change the direction and methods of subsequent investigations. We will begin with only a few subjects. If results are encouraging, later studies will involve more subjects. If initial studies are successful, classroom teachers will be taught how to include self-control skills in their curricula. It is also possible that parents and others may participate.

Ideally, this research approach will yield information on which self-control skills can be taught and how to teach them to severely handicapped individuals. We hope that the product of UWRO's research in self-control will be a package of materials for trainers to use in teaching self-control skills to severely handicapped individuals in school, home, community, and vocational settings. Whatever the results, we expect that research in self-control will produce information vital to our understanding of and training for generalization.

Summary of First Year's Findings

The first year's self-control studies were designed to investigate questions relating to self-monitoring: if and how self-monitoring skills can be acquired by severely handicapped students, and the effect of self-monitoring on skills in training and generalization settings. Three studies were conducted by Dr. Kathleen Liberty; in two studies, extensive cooperation was provided by Mary Ann Paeth of Central School District 13J in Independence, Oregon. For more information on these studies, see Liberty (1983) and Liberty & Paeth (1983a, 1983b).

Self-Monitoring 1-1: The acquisition of self-monitoring and its effect on the production rate of a severely handicapped adolescent. The purpose of the first study was to examine (1) the acquisition of self-monitoring by a severely handicapped student through an avoidance training procedure and (2) the effects of self-monitoring on the target behavior.

The subject was 19.5 years old (IQ of 35 and MA of 4.5 years). During the training phases, the subject was taught to push the plunger of a counter placed on the table next to his work, using an avoidance training procedure. In the second training phase, the subject wore the counter on his wrist. In the last phase, a general contingency for tantrumming was introduced, and was in effect throughout the school day.

The avoidance training procedure produced rapid acceleration of independent use of the counter without the addition of specific reinforcement for self-monitoring. Very high levels of reliability were obtained without specific reinforcement for reliability, and without specific cues directed at pushing the counter only one time. However, since perfectly independent and reliable self-monitoring was not produced until indiscriminable contingencies were introduced, training in the use of a counter may be accelerated by providing reinforcement specifically for independent and reliable self-monitoring.

Data on the target behavior indicate that self-monitoring acted as a positive reinforcer sufficient to maintain performance of the target behavior. Thus, self-monitoring itself is reinforced by the counter and the act of self-monitoring functions as a self-controlled positive reinforcer for production. Overall, however, the changes in production during the course of the study are unlikely to have any practical impact on the subject, both because overall improvement did not move the subject significantly closer to a normal work rate, and because the work itself is not performed outside of classroom settings.

Self-Monitoring 1-2: Maintenance and generalization of self-monitoring and its effects on two target behaviors. The purposes of this study were to (1) examine the maintenance of self-monitoring skills; (2) determine if self-monitoring generalized within stimulus classes and across responses, and if so, the extent and nature of such generalization; and (3) examine the effect of self-monitoring on the target behaviors. The subject was the same subject who participated in the first self-monitoring study.

Self-monitoring maintained at high levels of reliability and independence. The data also provide a measure of the generalization of self-monitoring across different stimulus conditions within the same class. When the stimulus materials were changed, reliable self-monitoring was established in one or two sessions. This level of maintenance and generalization may be interpreted as of practical value, since instruction time is "saved," but its ultimate value is questionable, since some level of avoidance prompting was required.

Independent and reliable self-monitoring did not transfer across behaviors. The subject did not actuate the counter when he had the opportunity to do so. In this study, the two behaviors were not of the same response class, and were performed under conditions (e.g., supervision, stimulus materials, setting, time of day) totally unlike one another. The failure to transfer presents a significant challenge to the training procedure used. In future studies, methods of changing the training procedure to enhance the probability of transfer will be examined.

When the self-monitoring avoidance prompting procedure was introduced for assembling sack lunches, a few prompts were sufficient to produce generalized counting of the next ten sacks at 100% reliability. However, avoidance prompting procedures were not sufficient to maintain counting throughout the period. It was only when a variable schedule of avoidance prompting was instituted that reliable and independent self-monitoring was maintained throughout the work period.

The most powerful effect on bagging was produced during the second phase, when the subject had the opportunity to self-monitor, although he was not observed to do so. Wearing the wrist counter produced an increase in rate that is of practical significance to the subject, whose median production rate increased to 67.5% of normal, and reached 114% of the normal rate on his best day.

Once the subject was trained to actually self-monitor, the highest production rates were associated with days of perfect reliability of the rated sacks, even though the self-monitoring added a movement to be completed. However, once the opportunity to self-monitor was removed from production, bagging rate dropped. This also suggests that self-monitoring mediated some of the differences between the settings. When the opportunity to self-monitor bagging was withdrawn, bagging began to decelerate.

Self-Monitoring 1-3: Effects of self-monitoring training on expressive communication: Mediation across settings. The purpose of this study was to extend and replicate results of Self-Monitoring 1-2, with different behaviors and with a different subject. Specific issues examined included (1) the effects of wearing a wrist counter on two behaviors, one instructed and the other uninstructed, prior to any training in the use of a counter, and whether effects produced in the training setting would transfer to the probe setting; (2) the effects on the instructed and uninstructed behaviors of training self-monitoring of the instructed behavior, in both the training setting and the probe setting; (3) whether self-monitoring would transfer from the behavior

on which it was trained (instructed behavior) to another behavior (uninstructed behavior); and (4) whether self-monitoring would transfer from the setting in which it was trained to the probe setting.

The subject of this study was an 11 year-old girl, attending a special school for handicapped children. School records indicate that IQ testing had never yielded a valid score, but psychologists' estimates were of an IQ between 30 and 32.

Independent variables included an instructed response and an uninstructed response, measured in both a training setting and in a probe (i.e., generalization) setting. The instructed response was two-word answers to questions (instructed in the training setting only). The uninstructed response was two-word initiations. Data were also collected on actuation of a wrist counter, and the independence and reliability of self-monitoring.

Despite acceleration of two-word answers in the training setting, generalization of the instructed response to the probe setting was low, and slightly decelerating during the phase, with a median level of 18%, and ending at 15%. Instruction in self-monitoring of two-word answers in the training setting produced rapid acceleration of independent self-monitoring.

Self-monitoring may not have affected any change in the instructed behavior in the training setting; however, the opportunity to wear the counter in the probe session resulted in an immediate and sustained change in the target behavior. The median level of two-word answers in this phase is 43%, more than twice the level of generalization in the previous phase. The subject was observed to actuate the counter in the probe setting. When the opportunity to self-monitor in the probe setting was withdrawn, performance decelerated to an ending level of 19%, comparable to performance in the first baseline. In addition, when the subject was given the opportunity to wear the counter during the final probe setting session, performance was comparable to that of the previous phase. Both the instructed and uninstructed target behaviors generalized to three- and four-word utterances. However, such generalization did not occur in the probe setting until self-monitoring phases.

These results support and expand the conclusions of the earlier two studies. By itself, the wrist counter does not substantially affect responding. However, once training in use of the wrist counter is initiated, the counter itself is able to mediate the differences in consequation between settings. Although the level of generalization of two-word answers mediated by the counter is a substantial improvement over nonmediated responding, the level of generalization attained is likely to be insufficient to result in practical improvement for the subject. However, training self-monitoring does seem to be a simple and somewhat efficient way to mitigate differences in settings when skill transfer is desired.

Summary of Second Year's Findings

The second year's self-control studies were designed (1) to continue the investigation of questions related to self-monitoring and (2) to extend these

investigations to include self-reinforcement procedures. Two studies were conducted by Dr. Kathleen Liberty and two research assistants with subjects in two local public school districts. For more information on these studies, see Liberty (1984a, 1984b).

Study 2-1: Effects of a generalization package with and without self-monitoring. This study was designed to compare generalized responding of profoundly handicapped students under two different training conditions, one in which the major recommendations for modifying instruction were integrated in an on-going instructional program as a "generalization package," and a second training condition in which instruction in self-monitoring was added to the package.

Three students attending a public school program for severely handicapped students served as subjects. Each of the subjects began receiving instruction in answering questions that required either a "yes" or a "no" response (hereafter called yes/no questions) three weeks prior to the start of the study. The yes/no response has an obvious face validity as a functional behavior which will access a natural community of reinforcers, one of the recommended ways of facilitating generalization (Stokes & Baer, 1977).

Commercially available counters were modified to permit actuation by the subjects. The counters were then mounted on the tray tables of the subjects' wheelchairs so that their functional range of motion would permit physical actuation of the counter and sight of the counter display. Handwritten signs, introducing the subject and her/his method of answering questions to nontrainers, were taped to the subjects' tables.

The instructional procedures were designed to exemplify the major programmatic modifications described by Stokes and Baer (1977). Two major dimensions of generalization were included in the study: (a) nontrainers as questioners in the training setting (i.e., generalization within setting across individuals) and (b) nontrainers as questioners across all nontraining settings (i.e., generalization across settings and across individuals).

Discussion. The results suggest that the opportunity to self-monitor, when added to a generalization package, increased the number and the speed of answering nontrainer questions when compared to the generalization package alone. All of the subjects generalized answering yes/no across a continually changing universe of question content, nontrainer questioners, and settings. Subjects were observed to respond yes/no in the presence of the natural stimuli, but were not observed to respond yes/no at other occasions, another indication of appropriate generalization. The frequency and speed of generalized answering by Subjects 1 and 3 improved when they had the opportunity to self-monitor. When Subject 2 self-monitored, the frequency and speed of answering improved; however, self-monitoring occurred so infrequently that its effects appear to be of little practical significance.

It is clear that the "generalization setting" varied considerably for each of the subjects of this study. Nontrainers clearly differentiated between the subjects in regards to question content and question frequency. It is probable that the physical appearance of the subjects provoked this discrimination,

since this held true even at the time of the first interaction with a subject. The differences appear to be related to the degree of handicap of the subjects. That is, questions were asked less often with simpler contexts and much longer periods of waiting of the most profoundly handicapped subject, Subject 1, while the reverse was true of Subject 3.

In this study, self-monitoring increased the frequency and speed of answering when it was used by the subjects and therefore functioned as self-controlled reinforcement. The results suggest further that Subjects 2 and 3 may have used self-monitoring to mediate the effects of the consequation deficiencies in the generalization environment; certainly this may be regarded as exemplary "self-control."

The integration of the nonverbal training procedure for teaching self-monitoring with the generalization oriented instructional package for yes/no answers produced extremely rapid generalization of self-monitoring but very erratic acquisition and maintenance of independent and reliable self-monitoring. These data support the idea that generalization may occur concurrently with acquisition rather than as a product of acquisition training. They may also indicate that some procedures (e.g., verbal prompts) most efficient for establishing strict stimulus control over the target behavior (acquisition), may be in conflict with training procedures that establish the "loose" stimulus control which appears necessary for generalization of certain types of responses to occur.

This study indicates that even individuals regarded as the most profoundly handicapped can acquire and generalize self-monitoring, a skill hypothesized as requiring high levels of cognitive activity. They can use self-monitoring to improve the frequency and speed of answering new questions asked by new people they meet in new settings. This achievement shows the development of self-control as well as the generalization of answering that is meaningful in accessing future community environments.

Study 2-2: Self-monitoring and self-reinforcement and their effects on question answering. Self-reinforcement is one of the most frequently taught self-control skills. Self-control packages, in which self-reinforcement is one component, have been used to facilitate generalization across time, across responses made by an individual, and across settings. The purposes of this study were (1) to determine the effects of using a nonverbal method to teach subjects who had been taught to self-monitor to self-reinforce the target behavior and (2) to determine the effects of instruction in self-reinforcement on the generalization of the target behavior.

The two subjects attended a public school program for severely handicapped students. Subject 1 was 10.9 years old at the start of the study. His most recent evaluation indicated an intelligence level of 12-18 months. Subject 2 was 9.2 years old. In her most recent evaluation, she was evaluated as demonstrating a mental age of 1-1.5 years, with language between the 12-16 month range. Each of the subjects began receiving instruction in answering questions that required a yes/no response three weeks prior to the start of the study.

A golf counter designed to be attached to a key chain was modified for Subject 1. The 0 and 5 digits of the counter were covered with black dots. The subject wore the counter affixed by a small link chain to one of the belt loops on the left hand side of his waistband.

Subject 2's counter was a hand tally counter, which was mounted on her communication board so that the digit display and the actuator were above the surface of the board. The 0 and 5 digits of her counter were covered with pink dots.

Subject 1 had a small velcro wallet attached with his keychain and counter to his belt loops. Subject 2 had a small velcro purse attached by a ring and a small chain to her communication board. During trainer-controlled and self-controlled reinforcement phases, pennies were placed in a box, which was available to the subject throughout the school day (i.e., carried from setting to setting by the subject or by the trainer or data collectors).

The instructional procedures were designed to exemplify the major programmatic modifications described by Stokes and Baer (1977), with the exception of the use of pennies as reinforcers. Training procedures also included the expected natural contingencies and consequences. If subjects answered the question, trainers provided a natural consequence implied by the question content.

Two major dimensions of generalization were included in the study: (a) nontrainers as questioners in the training setting (i.e., generalization within setting across individuals) and (b) nontrainers as questioners across all nontraining settings (i.e., generalization across settings and across individuals).

Subject 1's acquisition of self-monitoring was rapid during the first self-monitoring phase. However, once training in self-reinforcement was implemented, Subject 1's independent self-monitoring decelerated. It improved once self-reinforcement training was withdrawn, but it did not accelerate.

Subject 2 did not demonstrate independent counting until the ninth session. Independent self-monitoring rapidly accelerated following that point and maintained at high levels until the end of the second self-monitoring phase.

Subject 1 had a total of 63 training trials in self-reinforcement. Of these, he independently self-delivered a penny when the counter was on black on only one occasion. He self-reinforced after counting on two other occasions when the counter was not on black.

Subject 2 had 30 self-reinforcement training opportunities. She self-reinforced independently twice when the counter was on pink, and three times when the counter was not on pink.

Subject 1 generalized self-monitoring to answers to nontrainer questions very infrequently throughout the study. Because of the few generalized self-

monitored answers, the counter never signaled a self-reinforcement opportunity with nontrainers for Subject 1. He was not observed to deliver a penny to himself for any response to a nontrainer.

Subject 2 generalized self-monitoring for the final three days of the first self-monitoring training condition, after 13 training sessions, and continued counting during each session of Phases 3 and 4. The average percentage of answers self-monitored was 45%. Because of the higher number of self-monitored answers, Subject 2 had a total of 14 opportunities to generalize self-reinforcement to answers to nontrainer questions. She did not deliver a penny to herself during any of these opportunities.

In general, the results suggest that the opportunity to self-monitor, when added to a generalization package, improved the level of generalized question answering for both subjects. The decelerating trends in the final phase, almost identical as the first phase for both subjects, indicate that the interventions did improve generalized answering overall. The interventions produced mixed effects on the speed of answering. Generally, Subject 1's speed improved while Subject 2's speed worsened when the self-control training procedures were in effect.

Probably the single biggest problem with the self-reinforcement training in this study was the paucity of training opportunities. This problem was caused by the selection of a FR5 schedule for token reinforcement. Since the subjects did not learn to self-reinforce, the effects of self-reinforcement cannot be evaluated in this study. Additional training opportunities could have been provided by beginning with a FR1 schedule. However, since behaviors reinforced under this type of schedule are subject to rapid extinction if the reinforcer is not available, it was decided to begin with a FR5 schedule. In order to avoid such a problem, it may be necessary to begin with FR1 to train self-delivery, and then to lean the schedule once self-reinforcement behaviors have been acquired.

Studies in Strategy Implementation

Meaningful instruction of severely handicapped pupils must insure that skills learned in specific environments (e.g., classrooms) will be used in other environments and under different stimulus conditions when such skills are required. Although an instructional technology is emerging which should facilitate generalization, skills taught to severely handicapped pupils in public school programs and in other nonlaboratory settings frequently fail to generalize. As a result, the capacity of severely handicapped person for independent functioning is often restricted.

The strategies included in the retrospective analyses, the ecological variables identified in the baseline studies, and the intervention strategies investigated directly by UWRO investigators have been implemented under typical experimental conditions in applied settings. The effects of such strategies or of changing certain ecological conditions when experimental controls are absent are unknown. Therefore, in addition to the continuing

experimental investigations, UWRO will conduct a major series of Strategy Implementation Studies, under the direction of Senior Investigator Valerie Lynch and Research Associate Frances McCarty.

The purpose of this series of studies is to test the effects of generalization-promoting strategies on skill generalization by severely handicapped students when such strategies are implemented by public school teachers under typical classroom conditions. The Strategy Implementation Studies will be conducted during the third and fourth project years.

Design of Intervention Studies

The procedural plan for this series of studies includes the following steps:

(1) Cooperating school districts will identify teachers and their severely handicapped pupils, following informed consent procedures approved by the University of Washington Human Subjects Review Committee.

(2) Teachers and UWRO staff will identify a behavior from each pupil's IEP, and settings where generalization of the behavior is both desired and appropriate. Although it is impossible to determine ahead of time what the goals of any specific subject will be, it is possible to provide some general guidelines. Behaviors chosen will be acceleration target behaviors—desirable and appropriate behaviors. Such behaviors may include self-care skills, like dressing and eating, communication skills, or academic tool skills. In addition, the teacher and researchers will identify one or more settings in which generalized responding is desired. The exact generalization setting is dependent on the target behavior. For example, the generalization setting for toileting will be bathrooms other than those bathrooms in which training occurs; the generalization setting for eating skills may be a restaurant in the community or the family kitchen if eating skills are trained in the school cafeteria; the generalization setting may be the school hallways and playgrounds for communication skills trained in the speech therapy room.

(3) The teacher will describe the instructional strategies currently implemented for the target behavior.

(4) Accuracy and fluency data on the performance of the target behavior for each pupil in one or more generalization settings will be collected ("Ongoing Instruction Phase"). An observation method will be used to check the use of the instructional strategies identified by the teacher.

(5) The classroom teachers will be trained by UWRO staff in (a) changing an ecological variable (e.g., number of settings accessed by students), (b) changing an instructional procedure currently in use (e.g., changing from artificial stimuli to common stimuli), (c) implementing a different instructional procedure (e.g., teaching self-monitoring), and/or (d) some combination of strategies. Teachers will be trained both in the general application, and in the specific application to their student(s).

(6) Accuracy and fluency data on the performance of the target behavior will continue to be collected during the implementation of the trained strategy by the classroom teacher ("Strategy Implementation Phase"). Project staff will observe to check the implementation of the strategy by the teacher.

The data on the generalized performance of the pupils will constitute the primary evaluative tool for assessing the effectiveness of identified strategies. The overall experimental design for the studies will be a simple AB design replicated across students and teachers within the same school.

Expected Results and Outcomes

The data collected as part of the Strategy Implementation Studies will ensure that procedures recommended for use by teachers on the basis of UWRO research will be ones which retain their effectiveness when applied under normal public school constraints and conditions. These data will provide the integral underpinnings of the recommended methodology for teachers that will be the most important product of UWRO activities.

References

Billingsley, F., Berman, A., & Opalski, C. (1983). Generalization and the educational ecology of severely handicapped learners: A descriptive study. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 82-83* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Billingsley, F., Thompson, M., Matlock, B., & Work, J. (1984). Generalization and the educational ecology of severely handicapped learners: A descriptive study—Second Year. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 83-84* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Kayser, J. (1984). The effects of two instructional strategies on the acquisition and generalization of a functional task with severely behavior disordered children: A brief report. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 83-84* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A. (1983). Self-monitoring III: Effects of s-m training on expressive communication: Mediation across settings. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 82-83* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A. (1984a). Effects of a generalization package alone and with self-monitoring on generalized yes/no answers of profoundly handicapped subjects. In N. Haring (Principal Investigator), *Institute for education of severely handicapped*

children: Washington Research Organization; Annual report FY 83-84 (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A. (1984b). Acquisition and generalization of self-monitoring and self-reinforcement and their effects on question answering. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 83-84*(U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A., & Paeth, M. A. (1983a). Self-monitoring I: The acquisition of self-monitoring and its effect on the production rate of a severely handicapped adolescent. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 82-83* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A., & Paeth, M. A. (1983b). Self-monitoring II: Maintenance and generalization of self-monitoring and its effects on two target behaviors. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 82-83* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Weisenstein, G., Field, S., & Kiolet, C. (1984). Investigating strategies for the systematic fading of cues, reinforcers, and contingencies. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 83-84* (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

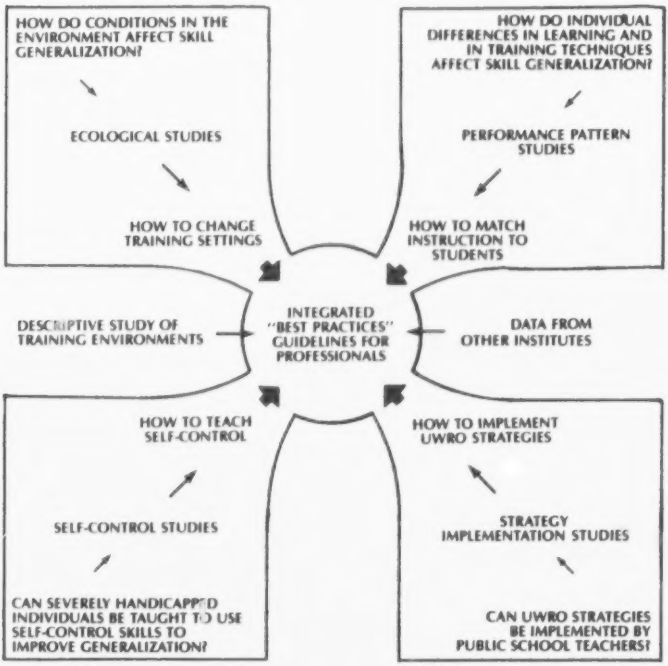
White, O. R., Haring, N. G., & Miller, S. B. (1983). Retrospective analysis of existing data sets: Performance patterns and instructional strategies for promoting skill generalization. In N. Haring (Principal Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual report FY 82-83*(U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

FOUR

PUTTING IT ALL TOGETHER

Guidelines for identification and manipulation of a wide range of conditions within educational settings will result from the studies in the ecology of training settings. The performance pattern studies will contribute a set of guidelines specifically for instructional methods that educators can use to ensure generalization. Guidelines from the ecology studies will be directed at fairly global management of the instructional setting, while decision rules from the performance pattern studies will be directed at the selection of precise instructional methods used in individual programs.

During the fourth project year, methods will be developed to combine the guidelines from the ecology and performance pattern studies with other empirical data, into an integrated set of "best practices for generalization."


HOW DO CONDITIONS IN THE ENVIRONMENT AFFECT SKILL GENERALIZATION?
ECOLOGICAL STUDIES
HOW TO CHANGE TRAINING SETTINGS
HOW DO INDIVIDUAL DIFFERENCES IN LEARNING AND IN TRAINING TECHNIQUES AFFECT SKILL GENERALIZATION?
PERFORMANCE PATTERN STUDIES
HOW TO MATCH INSTRUCTION TO STUDENTS
DESCRIPTIVE STUDY OF TRAINING ENVIRONMENTS
INTEGRATED "BEST PRACTICES" GUIDELINES FOR PROFESSIONALS
DATA FROM OTHER INSTITUTES
HOW TO TEACH SELF-CONTROL
SELF-CONTROL STUDIES
CAN SEVERELY HANDICAPPED INDIVIDUALS BE TAUGHT TO USE SELF-CONTROL SKILLS TO IMPROVE GENERALIZATION?
HOW TO IMPLEMENT UWRO STRATEGIES
STRATEGY IMPLEMENTATION STUDIES
CAN UWRO STRATEGIES BE IMPLEMENTED BY PUBLIC SCHOOL TEACHERS?

Such guidelines would probably establish a decision hierarchy for use at administrative, training setting, small group, and individual pupil levels. For example, a sequence of decisions might include:

(1) Determine what skills should be programmed for generalization.
(2) Determine the appropriate instructional settings (i.e., home, school, or community) for each skill.
(3) Determine the characteristics of the setting in which generalization is desired.
(4) Determine for each student the percentage of each school day to be spent in each setting, or how to integrate factors from the generalized setting into the training setting
(5) Determine if instructional trials will be massed or distributed.
(6) Determine the specific instructional procedures for each student.

The guidelines that may result from the studies in self-control will affect the curricula of training settings by suggesting changes in the skills that are currently taught. Recommendations, such as the inclusion of self-monitoring in the curriculum, may be accompanied by precise directions as to whom to teach such skills and how such skills might be most effectively taught. It is expected that information on other curricular changes that affect generalization, produced by the work of other Institutes, would be used to produce a set of integrated guidelines for curriculum content. If all conditions are ideal, perhaps the guidelines for curricula will be integrated with the guidelines for intervention in the setting and included as an aspect of the data decision rules, producing a fully integrated single set of practices.

At this time it is difficult to predict the nature of the various guidelines to be developed or if the guidelines will fit together, since they must be based on empirical evidence that the strategies do, in fact, promote generalized responding. The strategy implementation studies will provide evidence of the effectiveness of UWRO's recommended generalization strategies. UWRO will draw on the expertise of the Advisory Committee and the Direct Service Consortium in the development of guidelines. We will also have access to results from the other Institutes. All of the information available will be integrated into the guidelines eventually produced. It is our hope that the four approaches will provide solutions converging into an integrated set of guidelines for users. The schematic, shown on page 58, illustrates how UWRO hopes to increase interaction and integration of the results as research proceeds, to the development of an integrated set of guidelines for practitioners.

The Washington Research Organization combines four different and complimentary conceptual approaches to the problem of skill generalization. We believe that pursuit of these lines of inquiry represents a strategy with the highest probability of defining replacements for the "train and hope" methods on which educators must currently rely. Implementation of the concept of a free appropriate public education in the least restrictive environment for all students should not be undermined by ignorance. The contributions the Washington Research Organization makes to the development of a technology of skill generalization are contributions to the work of all who strive for the realization of our social commitment to an effective and lasting education for all severely handicapped individuals.

Cindy is apprehensive her first day on the job at the Pacific Oyster Bar. She failed so badly at the Seattle Hotel. She looks carefully at the dishwasher, and loads the bowls and cups. She closes the door. She searches and finds the buttons on the side of the machine. They are strange, but the little stickers just below them are just like the ones at school. She confidently pushes the series, and smiles when the dishwasher hums into action. At the end of the day, the kitchen supervisor says, "Good work today, Ms. Burchart." He smiles as Cindy gets her coat and leaves. Still smiling, he looks again at the little stickers the trainer from the Seattle Training Center had put on each of the dishwashers. He thinks, "Well, you learn something new every day."

Richard leaves the office of the head housekeeper. As he wheels himself toward the chain of pink cabins of the Sunset Motel, he repeats to himself, "Knock. Then say, 'Housekeeping here.'" Over and over he says these instructions, just as Mr. White taught him to do when he was teaching him to say his name and address, all those years ago. He is pleased that he can practice by himself. At Cabin 1 he stops, squares his shoulders, and knocks briskly. "Housekeeping here." He unlocks the door and goes in to earn his first wage.

Jody is screaming so loudly that his face is eggplant purple again. Mrs. Loomis smiles to herself, and walks out the door to join the rest of the family waiting in the car, leaving Jody's jacket on the floor where he threw it. She gets in the car. "Now where's Jody?" asks Mr. Loomis. "Just wait," she replies. In 30 seconds, Jody comes flying out the door, zipping his jacket. "Don't forget to shut the door," cries his mother. She thinks with satisfaction of Jody's teacher—she was right, after all! Jody does know how to put on his jacket.

Part 2:

Literature Reviews

FIVE

TRAINING IN THE NATURAL ENVIRONMENT AND SKILL GENERALIZATION: IT DOESN'T ALWAYS COME NATURALLY

Owen R. White
B. Douglas Leber
and Claire E. Phifer

In the search for strategies which will facilitate the development of community integration skills, "training in the natural environment" has gained considerable popularity. Although the bulk of research to date has investigated relatively limited applications of training in the natural environment, comprehensive systems for school-home-community training programs have been proposed (cf., Sailor & Guess, 1983) and it has even been suggested that all training might best be undertaken within the community at large (cf., Brown, Nisbet, Ford, Sweet, Shiraga, & Loomis, 1982). After reviewing a number of approaches to community training, Sailor & Guess concluded that:

> The future of elementary programs for the severely handicapped will clearly lie in a school-without-walls concept, which, although based on the elementary school campus, will focus increasing amounts of time off-campus in a multiplicity of environments. (p. 274)

Somewhat later in their analysis, Sailor & Guess also conclude that:

> [Total] nonschool instruction is in its early developmental period . . . In our opinion, however, the model is completely realistic and worthy of large-scale research efforts. (p. 318)

Without doubt, the movement toward community-based training is rapidly gaining momentum. Such a movement might be justified as a means of achieving integration in its own right by providing an opportunity for interaction among handicapped and nonhandicapped persons. Skills acquired in the natural environment are also of at least limited functionality without the need for further generalization. The validity of those arguments will not be debated here. It will be argued, however, that regardless of the immediate functionality of a skill, it is still desirable for a training program to facilitate appropriate skill generalization across as broad a range of meaningful environmental variations as possible. The purpose of the study presented here was to examine the implications of training in the natural environment with respect to that issue.

Procedures

A systematic review was undertaken of the research literature pertaining to training strategies which might be appropriate for use with persons experiencing severe handicaps. A total of 11 journals were examined[1] covering the years 1977 to 1984.

An emphasis was placed on the review of articles involving severely handicapped subjects. For comparative purposes, however, articles were also included which studied mildly, moderately, or nonhandicapped subjects.

Articles were included in the review only if they presented original research data concerning the performances of one or more individual subjects. For reasons which will become apparent later in this paper, articles presenting only group-level summary data were not included.

All reviewed articles presented data concerning skill generalization of at least one functional behavior across at least one meaningful stimulus/environmental dimension. Consensus among reviewers regarding what constituted "functional behavior" and "meaningful stimulus/environmental dimensions" was in question with circa 10% of the studies reviewed. However, subsequent comparisons of the outcomes of those studies with the outcomes of studies dealing with clearly functional behaviors and conditions of generalization failed to show any systematic differences, so for purposes of the questions addressed in this article, no further distinction among those cases will be made.

A total of 115 studies were found which met the conditions outlined above[2]. A brief overview of the characteristics of those articles may be found in Table 5-1.

Coding Form

Information concerning each article was coded using a form developed by the Washington Research Organization[3]. The form allowed 145 pieces of information to be recorded for each article, including article identification, subject characteristics, target behavior types, training and generalization conditions, generalization-facilitation strategies (if any), performance data types, initial subject performance levels, and the degree to which the target behavior generalized from training to nontraining situations. The investiga-

---

[1] Articles from 11 journals were included in the retrospective analysis. Circa 48% of all articles were published in the *Journal of Applied Behavior Analysis*, 13% in the *Journal of the Association for Persons with Severe Handicaps*, 9% in *Behavior Modification*, and 8% in *Education and Treatment of Children*. The remaining 22% of the articles were distributed among the *American Journal of Mental Deficiency*, *Behavior Therapy*, *Behavior Research and Therapy*, *Education and Training of the Mentally Retarded*, *Journal of Experimental Psychiatry*, *Mental Retardation*, and the *Journal of Applied Research in Mental Retardation*.

[2] A complete list of review articles may be obtained from the authors upon request.

[3] A copy of the recording form, the coding manual, and/or computer listings of actual codes may be obtained from the authors for cost of reproduction and mailing.

Table 5-1
Characteristics of Reviewed Articles

	Not Severely Handicapped	Severely Handicapped	TOTAL
Total Number of Studies	44	71	115
Total Number of Subjects	156	249	405
Subjects/Study (range)	1-12	1-10	1-12
(median)	3	3	3
(mean)	3.5	3.5	3.5
Subjects Age in Years (range)	3-52	<1-63	<1-63
(median)	10.5	14.0	12.0
(mean)	15.8	15.1	15.4
"Natural Setting"			
(studies)	52.3%	45.1%	47.8%
(subjects)	57.1%	42.2%	47.9%

tion presented here focused only on information concerning basic subject type (severely handicapped vs. all other subjects), the utilization of basic generalization-facilitation strategies, and the level of generalization achieved.

Subject-Type Codes

Subject type was originally coded as one or more of the following: nonhandicapped, mildly mentally retarded, moderately mentally retarded, severely mentally retarded, profoundly mentally retarded, multiply handicapped, physically handicapped, visually impaired, auditorily impaired, communication disabled/impaired, severely behaviorally disturbed/autistic, or deaf-blind. For purposes of the current investigation, subjects were reclassified as being either "severely handicapped" (i.e., severely mentally retarded, profoundly mentally retarded, moderately mentally retarded with severe physical disabilities, multiply handicapped, severely behaviorally disturbed/autistic, or deaf-blind) or "not severely handicapped" (i.e., all other categories). The occurrence reliability of 6 independent raters on a sampling of 12 (10.4%) of the articles ranged from a low of 50% to a high of 100% when considering each subclassification of subject type. Overall reliability on coding major subject types (severely handicapped vs. not severely handicapped) was 100%.

Generalization Strategy Codes

Codes were developed to record any generalization-facilitation strategies employed in each study. The list of possible strategies was based originally on those suggested by Stokes & Baer (1977), but was expanded and modified somewhat during the course of the study to account for new developments in the field and to allow finer discriminations among strategy types. Each of the strategies is defined briefly with examples in Table 5-2.

Table 5-2
Training Strategies for Promoting Generalization[1]

	STRATEGY AND DEFINITION	EXAMPLE
Class A[2]	**Program Common Stimuli:** Salient stimuli from target setting are introduced into training setting. Such stimuli may be irrelevant to target behavior but still control generalization.	In teaching group P.E. skills, music used in the group class (target setting) is used in individual training as well (Stainback et al., 1983).
	Train Loosely: Task-irrelevant stimuli are intentionally incorporated and varied in the training procedures. Neither the stimuli nor the response required in skill acquisition are allowed to be ritualized. A topographic variety of responses may be accepted as correct.	Training a language response (yes/no reversal, "wh questions") is conducted concurrently with training other skills. Child could respond to variety of natural events or teacher "set-up" events to evoke a response (Campbell & Stremel-Campbell, 1982).
	Introduce to Natural Maintaining Contingencies: Control of the target is transferred to stable, naturally occurring communities of reinforcement. The authors present evidence that the natural contingencies are reinforcing for the subjects, and the subject is allowed to "sample" the reinforcement.	Cafeteria-style meal service is replaced by family-style meals, promoting meal-time language use (VanBiervliet et al., 1981). Children are taught to solicit praise for their good work (Stokes et al., 1978).
Class B[3]	**Sequential Modification:** Training is conducted in one setting and generalization is probed in one or more non-training settings. If generalization is not evident, training is introduced sequentially to the probe settings. This process is continued until generalization is achieved in at least one non-training setting or training has been completed in all target settings.	In one setting, subjects were trained to articulate a word in response to a picture of it. Generalization was probed in 4 other settings. Training was introduced into one probe setting (Murdock et al., 1977).
	Sufficient Exemplars: Stimulus exemplars are trained sequentially in training setting until generalization to non-training exemplars is observed or all target exemplars have been trained.	Subjects were taught to use self-instruction to focus attention and cope with 2 academic tasks: math and printing. Several distracting situations were introduced sequentially: photo-slides, audio-distractors, and kindergarten children playing with wooden blocks in the training setting. Increases in attending behavior were observed in non-trained academic programs. (Burgio et al., 1980).
Class C[4]	**Multiple Exemplars:** Several exemplars are concurrently used in training. No systematic attempt is made to represent the range of stimulus variations of the response class.	Three different machines were used to teach vending machine use. (Sprague & Horner, 1984).
	General Case: Exemplars chosen for training are systematically selected to represent the range of stimuli included in the category or response class.	In contrast to the multiple exemplar approach, in choosing vending machines for use in training, a number of machines were analyzed and three examples were chosen which, among them, represented the range of stimulus variations likely to be encountered by the subjects (Sprague & Horner, 1984).
	Train and Hope: This strategy is defined by default. That is, a study was classified as "Train and Hope" if there was no explicit programming for generalization or if procedures were used which would not be expected to facilitate generalization.	

[1]Several other strategies for promoting generalization were also coded. These are the strategies which Stokes and Baer (1977) characterized as Training to Generalize, Introduce Indiscriminable Contingencies, and Mediate Generalization. Too few articles were coded as using these strategies to include in the current study.
[2]These definitions are similar to those used by Stokes and Baer (1977).
[3]The following definitions are restricted examples of the categories introduced by Stokes and Baer.
[4]These strategies were not included in the Stokes and Baer article.

BEST COPY AVAILABLE

Several problems were encountered when attempting to code generalization strategies. First, many articles failed to provide sufficiently detailed information concerning the procedures they employed. That was especially important when making discriminations among various strategies which entail special pretraining preparation in materials or procedures. For example, one of the distinctions between "multiple exemplars" and "general case programing" is the level of analysis performed to ensure that exemplars adequately represent the "universe" of situations to which generalization is desired. When authors failed to provide information concerning the manner in which exemplars were selected, the code "multiple exemplars" was used, perhaps erroneously. Similarly, many articles may have been incorrectly coded as "train and hope" simply for lack of adequate procedural description.

Coders were occasionally misled by direct statements made by authors concerning the strategies they employed. In one study, for example, authors stated that all the behaviors they trained could be "viewed as those that are likely to be maintained by the natural consequences in a classroom" (Reese and Filipczak, 1980). That statement led at least one coder to select "introduce to the natural maintaining contingencies" as one of the strategies employed in the study. The authors went on to say, however, that "no systematic analysis of (nontraining classes) was conducted . . . " (p.221). In fact, no evidence was provided in the study to suggest that those natural consequences ever actually occurred or, if they did, that they were reinforcing to the subject.

Finally, even when procedural descriptions were reasonably precise, coders often had difficulty in deciding how those procedures should be coded. For example, several studies involved behavior which might be appropriate in virtually any environment (e.g., general social skills; Lancioni, 1982). Because the target behavior was so universally applicable, some coders recorded such studies as "train in the natural environment" regardless of how or where training actually took place. Other coders considered more carefully whether the training situation represented the range of environments in which the skill should be employed (e.g., different areas of the school, home, and community) and/or the general conditions under which training actually took place (e.g., massed trials at some arbitrary time of day versus distributed trials upon naturally occurring opportunities for the behavior). The latter approach was finally adopted as the standard, but only after discussion of many different examples. Eventually, it became clear that "natural" is a matter of degree, not just kind.

As difficulties with the coding system and criteria for classification were identified, coders would discuss the issues and decide upon revisions in the coding form. In all, seven successive versions of the coding form were developed. Overall coder reliability across all items on the original form was 75%, increasing to 86.8% on subsequent versions.

The reliability of six independent raters forming 21 comparisons of the strategy-coding section across a sampling of 13 studies (11% of all studies reviewed) was calculated two ways. In the first method, an agreement between coders was scored if both indicated a strategy had or had not been

employed. Reliabilities calculated via that method ranged from a low of 75% to a high of 100%, with a cross-study mean of 88.7%. Such general reliability statements can be inflated, however, if a large proportion of the coding opportunities are left blank (i.e., only one or two of the 12 possible strategies is coded for any given article). For example, if one coder indicated that a study employed only "train and hope" while the other coder indicated that the same study employed only "multiple exemplars," the overall agreement would still be 83.3%, simply because both *coders left ten of the twelve* possible strategies blank.

A more conservative method of calculating reliabilities is to consider only those strategies which one or both coders indicated were used in each study. Such "occurrence reliabilities" (Hawkins and Dotson, 1975) were much lower overall, ranging from 33% to 100% on individual studies, with a cross-study mean of 60.9%.

Since the discrimination of "training in the natural environment" was of special importance to the current investigation, an additional 23 studies (20% of the total number reviewed) were selected for reliability checks of just that code. Occurrence reliability for those studies was 74%. Subsequent review of the coding errors suggests that at least three of the six errors were equivocal, dealing with cases in which natural situations were at least simulated with reasonable fidelity (i.e, training social skills in a mock "free-play" situation with handicapped and nonhandicapped peers; Lancioni, 1982). If those high fidelity simulations were accepted as instances of a "natural setting," overall occurrence reliability for coding that strategy would be raised to circa 87%.

Given the conservative manner in which reliabilities were calculated, overall agreement on strategy codes was considered acceptable for purposes of the current investigation. However, it cannot be emphasized too strongly that even those modest levels of independent-coder reliabilities were achieved only after extensive discussion of earlier disagreements. Clearly, the identification of various strategies for the facilitation of skill generalization is a complex task. A more flexible, definitive classification schema should be developed.

Outcome Codes

The vast majority of studies concerning skill generalization with severely handicapped subjects employ a replicated single-subject time-series design. The outcomes of such studies are not easily summarized with simple statistics like means or standard deviations. At very least, one must consider baseline levels of performance, trends in baseline performance, the immediate effect of introducing the treatment, and changes in trend associated with treatment. Each of those performance descriptors must also be held relative to the variance or "bounce" in performance within phases of the experiment and, where possible, one must try to account for regular or irregular patterns in the data (e.g., autoregressions, integrated moving averages, seasonal cycles) which might influence interpretations of cross-phase performance

changes. Finally, the magnitude of the treatment effect must be held relative to some standard or criterion in order to determine whether any effect was "functionally meaningful."

Various descriptive and inferential statistics do exist to assist in the evaluation of each of those aspects of time-series designs. Since individual descriptors only deal with one aspect of the performance, however, they can drastically oversimplify what is often a very complex phenomenon, and it has been suggested that they only be used when the complete charted performance record is also available for inspection (White, 1984). Unfortunately, it would not be possible to present individual charts in a retrospective analysis of this type, so it was decided to rely upon the judgment of individual coders in evaluating study outcomes. Coders inspected progress records provided in the study and classified each subject as generalizing "well, some, or none."

In order to classify a subject as generalizing "well," two conditions had to be met. First, the subject must have demonstrated a functionally meaningful level of performance within a generalization setting. For example, a subject who generalized enough steps in a bus-riding sequence to independently ride a bus on an untrained route would be considered to have acquired a meaningful level of performance (Neef, Iwata, & Page, 1978). A subject who generalized all the steps for boarding a bus, but failed to generalize critical skills for determining when and how to get off the bus, would not be considered to have generalized well. Secondly, the study as a whole must provide convincing evidence that the behavior in the generalization setting was, in fact, generalization, and not the result of direct training in that setting. For example, a subject who required special prompts in a novel setting before beginning to engage in the target behavior would not be considered to have generalized well, even if only a few prompts were needed to produce independent performance at criterion level (Hill, Wehman, & Horst, 1982).

A subject was classified as generalizing "some" if increases in the performance of the subject in the generalization setting occurred which could reasonably be attributed to the treatment, but which failed to bring the subject to a completely independent, meaningful level of performance within the generalization setting without further training.

Subjects were classified as generalizing "none" if no increases in performance were noted in the generalization setting, or if such increases as were observed could not reasonably be attributed to training outside that setting.

For purposes of the current investigation, it was decided to concentrate on subjects who generalized "well" and achieved meaningful levels of performance without extra training in the target situation. In all of the analyses presented below, therefore, subjects who generalized "some" and "none" were combined into a single "not well" category. In a sampling of 21 ratings by six independent coders, overall agreement concerning major outcome (i.e., "well" versus "not well") was 80.9%.

Attributing Outcomes to Strategies

The overall purpose of the investigation was to determine how individual programming strategies appeared to facilitate skill generalization. Unfortunately, two problems prevented the direct assessment of the absolute effect of any given strategy.

First, the majority of reviewed articles had one or more design flaws which weakened their internal validity. The most common confound was some invariant sequence of treatment phases. For example, in order to assess the relative importance of using a general case training strategy, Sprague & Horner (1984) first trained subjects using a single exemplar to show that generalization did not occur. Unfortunately, even after generalization was obtained following the use of a general case approach, one could not determine whether it would have been obtained without initial training in the single case.

Secondly, aside from studies employing only a "train and hope" approach, only 31% (n = 31) studies employed a single, well-defined strategy to facilitate generalization. Even in cases where the experimental design provided good internal validity with respect to the effect of the "treatment package," therefore, it was often not possible to attribute effects to any given element of that package.

As a result of those problems, reviewers were instructed to simply code all strategies employed during the course of the study and to evaluate the overall level of generalization achieved. Basic issues of internal validity were taken into account as described earlier, but no attempt was made to ascribe effects to a single strategy where multiple strategies were used. The impact of that approach on the current investigation is likely to be a general "averaging of effects" across strategies. That is, less effective or noneffective strategies may still appear to be associated with some success to the extent that they were used in conjunction with more effective strategies. Similarly, the outcomes of highly effective strategies may have been attenuated through association with less effective strategies. When averaged across all studies reviewed, however, it is reasonable to expect the relative standings of strategies with respect to outcome to be meaningful. More effective strategies should be more consistently paired with better outcomes than less effective strategies. Still, it is important for the reader to bear in mind that the actual number or proportion of successful outcomes listed for each strategy may be somewhat inflated or deflated through association with other strategies.

Results

Major Training Conditions

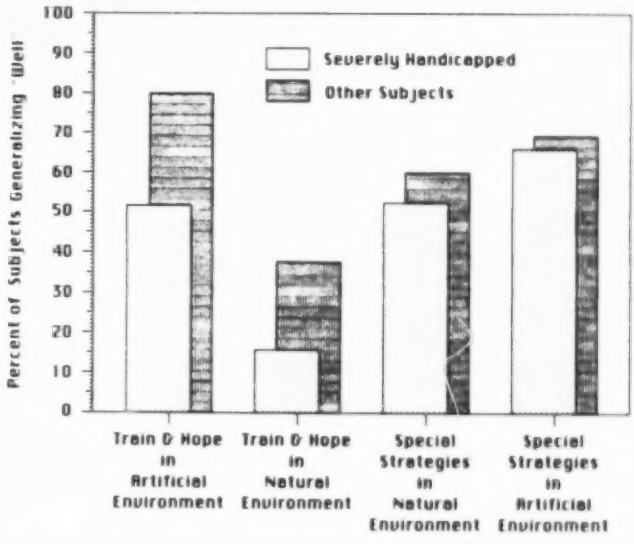
Table 5-3 and Figure 5-1 summarize the proportion of subjects across all studies generalizing "well" following training under each of the major conditions studied: natural and artificial settings, with and without the use of special strategies designed to facilitate generalization.

Table 5-3
Percent of Subjects Generalizing "Well" Following Training in Artificial and Natural Environments

	Train and Hope in Artificial Setting	Train and Hope in Natural Setting	Special Strategies Natural Setting	Special Strategies Artificial Setting	Total
Severe Subjects	17/33 51.5%	3/19 15.8%	45/86 52.3%	73/111 65.8%	138/249 55.4%
Other Subjects	12/15 80.0%	11/29 37.9%	36/60 60.0%	36/52 69.2%	95/156 60.9%
Overall	29/48 60.4%	14/48 29.2%	81/146 55.5%	109/163 66.9%	233/405 57.5%

Note. Cell contents: The number generalizing "well"/total in sample is displayed above the percent generalizing "well."
Severely handicapped: $x2 = 17.424$, 3df, $p < 0.001$.
Other Subjects: $x2 = 10.259$, 3df, $p < 0.025$.

Figure 5-1
Percent of Subjects Generalizing "Well" Following Training in Natural and Artificial Environments



82

The greatest success with the severely handicapped followed the use of special strategies in an artificial environment (65.8% generalizing well). The lowest proportion of successful outcomes was associated with a train-and-hope strategy in the natural environment (15.8%). Train-and-hope training in artificial environments and the use of special strategies in the natural environment were both associated with moderate outcomes (51.5% and 52.3% generalizing well, respectively) and approximated the average for all severely handicapped subjects (55.4%).

The overall chi-square across categories of training for the severely handicapped was highly significant ($x2 = 17.424$ with 3 degrees of freedom, probability < 0.001). However, paired comparisons between training conditions reveal that all individually significant differences are associated with a train-and-hope training program in the natural environment. Significantly fewer severely handicapped persons generalize well following the use of such an approach than with any other of the approaches (see Table 5-4).

The pattern of results obtained for persons who were not severely handicapped is very similar to that obtained with the persons who were experiencing severe handicaps. Although higher proportions of the nonseverely handicapped generalized well under each condition (60.9%, overall, versus 55.4% for the severely handicapped), those differences did not prove to be statistically significant. A train-and-hope approach in the natural environment was associated with the least success (37.9% generalizing well), and the use of special strategies in natural and artificial environments were associated with progressively more successful outcomes (60% and 69.2% generalizing well, respectively).

Unlike the severely handicapped, the greatest success with those who were not severely handicapped was associated with a simple train-and-hope

Table 5-4
Chi-square Values and Probabilities for Differences in Proportion of Subjects Generalizing "Well" Following Various Training Approaches

	Train & Hope		Special Strategies	
	Artificial Setting	Natural Setting	Artificial Setting	Natural Setting
Train & Hope				
Artificial Setting		7.01 oth (0.01)	2.08 oth (not sig.)	0.66 oth (not sig.)
Natural Setting	6.50 sev (0.025)		3.82 oth (0.05)	7.49 oth (0.01)
Special Strategies				
Artificial Setting	0.01 sev (not sig.)	8.37 sev (0.005)		1.03 oth (not sig.)
Natural Setting	2.20 sev (not sig.)	16.69 sev (0.001)	3.64 sev (not sig.)	

Note. sev = chi-square for severely handicapped subjects
oth = chi-square for other subjects

approach in an artificial environment (80% generalizing well). That may be due in part to the nature of the behaviors investigated (frequently more "academic" in nature), or simply a reflection of the small number of subjects studied under that condition (n = 15). As with the severely handicapped, the overall chi-square for outcomes across training conditions for other subjects was highly significant ($x2 = 10.259$ with 3 degrees of freedom, probability < 0.025), and only the train-and-hope approach in the natural environment proved to be significantly different from each of the other strategies (see Table 5-4).

Specific Strategy Outcomes

Additional analyses were conducted to evaluate the proportion of subjects generalizing following training in each of the specific strategies studied. Due to the limited number of subjects available for study, it was not possible to evaluate all combinations of strategies. Subjects were simply classified as having been exposed to a particular strategy within an artificial training situation or within a natural training situation. It is important to note, therefore, that results across strategy-types are not independent of one another. Subjects exposed to more than one strategy were included in the analyses of all those strategies. As mentioned earlier, that interdependence may have a general averaging effect (raising the apparent success of less effective strategies; lowering the apparent success of more effective strategies), but standings among various strategies should still be indicative of relative effect, and within-strategy/cross-situation statistical comparisons are still valid.

Aside from "training in the natural environment" per se, data were collected on a total of 11 different strategies. Sufficient data were available for the analysis of the 7 of those strategies: general case programming, multiple exemplars, programming common stimuli, sequential modification, train and hope, introducing to natural contingencies, and the use of sufficient exemplars. While no studies were coded for the use of multiple exemplars with nonseverely handicapped subjects in the natural environment, data did exist concerning the use of that strategy with the severely handicapped, so it was retained in the analysis. Three of the remaining four strategies were dropped from the primary analysis because no data were coded concerning their use with severely handicapped subjects in the natural environment (i.e., mediate generalization, loose training, and train-to-generalize). The use of indiscriminable contingencies was also dropped from further consideration due to the small number of severely handicapped subjects for whom data were available (n = 3). The results associated with each strategy may be found in Table 5-5, and the results obtained with the 7 strategies for which comparative analyses were possible are illustrated in Figure 5-2.

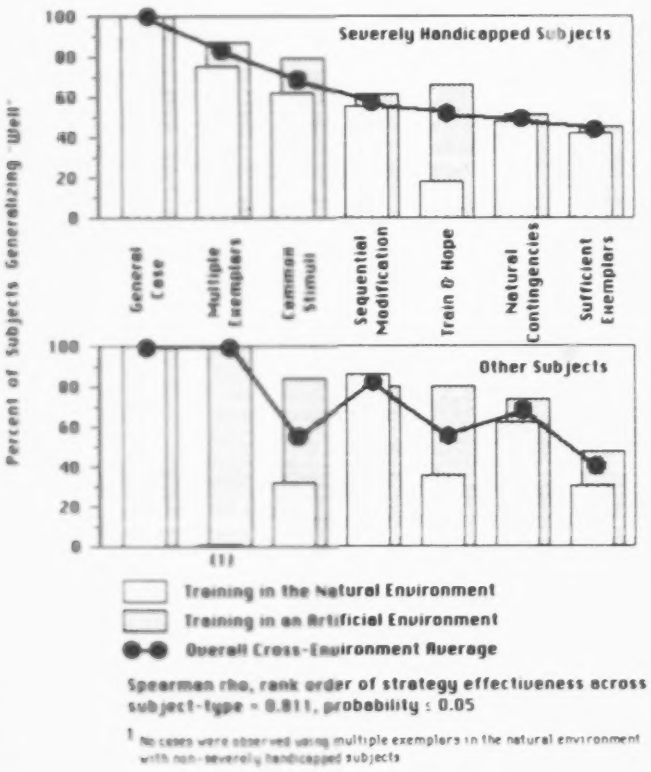
In all but one comparison, a greater proportion of subjects generalized well following training in an artificial environment than in a natural environment. In the exception to the rule, a higher proportion of nonseverely handicapped subjects did generalize well following sequential modification in the natural environment than in the artificial environment (85.7% versus 80%, respectively), but that difference is smaller than the record floor for

Table 5-5
Percent of Subjects Generalizing "Well" Following The Use of Various Special Strategies

	Severely Handicapped		Other Subjects		
	Artificial	Natural	Artificial	Natural	Total
(1) General Case Programming	13/13 100.0%	2/2 100.0%	5/5 100.0%	4/4 100.0%	24/24 100.0%
(2) Multiple Exemplars	14/16 87.5%	6/8 75.0%	3/3 100.0%	*	23/27 85.2%
(3) Common Stimuli	27/34 79.4%	30/48 62.5%	16/19 84.2%	8/25 32.0%	81/126 64.3%
(4) Sequential Modification	11/18 61.1%	21/38 55.3%	8/10 80.0%	6/7 85.7%	46/73 63.0%
(5) Train & Hope	43/65 66.2%	5/27 18.5%	20/25 80.0%	11/31 35.5%	79/148 53.4%
(6) Natural Contingencies	17/33 51.5%	28/58 48.3%	25/34 73.5%	25/40 62.5%	95/165 57.6%
(7) Sufficient Exemplars	10/22 45.5%	8/19 42.1%	8/17 47.1%	3/10 30.0%	29/68 42.7%
(8) Mediate Generalization	11/12 91.7%	*	0/5 0.0%	9/20 45.0%	20/37 54.1%
(9) Loose Training	15/20 75.0%	*	4/8 50.0%	17/17 100.0%	36/45 80.0%
(10) Indiscrim. Conting.	0/1 0.0%	2/2 100.0%	*	11/11 100.0%	13/14 92.9%
(11) Train to Generalize	2/13 15.4%	*	*	*	2/13 15.4%
(12) Train in Natural Setting	*	50/105 47.6%	*	47/89 52.8%	97/194 50.0%

*No cases coded in literature review

Figure 5-2
Percent Generalizing "Well" Following Training in Natural and Artificial Environments with Special Training Strategies



either condition (i.e., smaller than would result by increasing or decreasing the count in either condition by a single subject). Actually, despite the consistency with which exposure to training in artificial environments appears associated with superior results, most of the differences are very modest and statistically irrelevant. Significant chi-squares were obtained in only three cases: the use of common stimuli with nonseverely handicapped subjects ($x2 = 11.869$ w/1 df, $p < 0.001$); and the use of a simple train & hope approach with severely handicapped subjects ($x2 = 18.046$ w/1 df, $p < 0.001$) and with nonseverely handicapped subjects ($x2 = 11.097$ w/1 df, $p < 0.001$).

Due to possible overlap of subjects, direct comparisons of strategy effectiveness within subject-type would be tenuous. However, a comparison of rela-

tive effectiveness across subject type reveals a highly similar rank-ordering of the strategies (Spearman rho = .811, with a probability < 0.05). Thus, some cross-validation of relative effectiveness statements seems to have been achieved.

Other Results

Four strategies were not included in the analyses discussed above due to an insufficient number of severely handicapped subjects receiving training in the natural environment. In each case, however, interesting results were obtained.

Aside from the use of general-case programming, attempts to "mediate generalization" were associated with the highest proportion of severely handicapped subjects generalizing well (91.7%). Results obtained with subjects who were not severely handicapped met with much more moderate success, however.

"Loose training" was also associated with a high proportion of severely handicapped subjects generalizing well (75%). The lack of research concerning the use of that technique in the natural environment is particularly disappointing, since it proved uniformly effective in the natural environment with subjects who were not severely handicapped.

The use of "indiscriminable contingencies" was consistently associated with successful generalization in the natural environment with both types of subjects, but the number of severely handicapped subjects exposed to that strategy (n = 2) was simply too small to draw any meaningful conclusions.

Finally, attempts were made with 13 severely handicapped subjects to train generalized responding directly (i.e., reinforce instances of generalized behavior), but all such training occurred in artificial environments, and only two subjects (15.4%) generalized well. Those results provide at least modest support for the conclusions reached by Stokes and Baer (1977) that, "Common observation suggests that the method often fails, and that when it does succeed, little extrinsic reinforcement is offered as a consequence" (p. 362).

Discussion

A retrospective analysis of independent research studies requires that a number of compromises be made. First, it must be assumed that the characteristics of interest in such studies can be readily identified and reliably coded. Little difficulty was encountered in classifying studies according to subject type, at least within the major categories of "severely handicapped" and "not severely handicapped," but each of the other parameters of classification posed potentially serious problems. Identification of training methodology according to the schema proposed by Stokes and Baer (1977) was not as straight-forward as hoped, and modest levels of agreement were achieved only after many revisions to the code definitions. Even classifica-

tion according to general type of training situation (i.e., "natural" vs. "artificial") proved only somewhat easier, resulting in occurrence reliabilities of between 74% and 87%, depending upon how one wished to treat "high fidelity simulations" of natural situations.

The second major compromise involved the manner in which outcomes of the individual studies were described. While it would have been advantageous to quantify each outcome, no single method for the the statistical treatment of single-subject, time-series data is entirely satisfactory. Relying on the simple judgments of study reviewers to classify subjects as generalizing "well" or "not well" proved reasonably reliable but, of course, reduced the power of the analysis to detect differences among various strategies.

Finally, although some 405 subjects were available for study, the numbers associated with individual subject-type and strategy-type combinations were too small for any but the most global comparisons. That problem, compounded with the fact that the design of most studies made it impossible to attribute effects to any single strategy, restricted inferential analyses to broad comparisons of "train & hope" versus the use of one or more "formal strategies," and "artificial training environment" versus a "natural training environment."

Despite all those problems and compromises, one result appears quite robust—a train-and-hope approach in the natural environment is not associated with a high frequency of successful skill generalization. When exposed to a simple train-and-hope approach in the natural environment, only 15.8% of all severely handicapped subjects generalized well, and only 37.9% of all subjects who were not severely handicapped. In contrast, 51.5% of the severely handicapped subjects and 80% of the other subjects generalized well when exposed only to the train-and-hope approach in artificial environments. However, training in the natural environment per se does not necessarily result in poor skill generalization. When combined with the use of one or more other strategies specifically designed to facilitate generalization, the proportions of subjects successfully generalizing usually improved to levels comparable to those associated with similar training in artificial environments.

The specific reasons for the apparent failure of training in the natural environment to facilitate skill generalization cannot be determined from the current study. To the extent that skill generalization is related to initial skill acquisition, the outcomes may simply reflect the difficulty in arranging repeated practice within at least some natural settings (e.g., while boarding and exiting a mock bus can be practiced often, the frequency of riding actual city buses would be necessarily limited; Neef, Iwata, & Page, 1978). The strategies investigated in the current study tended to focus more on the selection and arrangement of stimuli rather than the frequency and duration of training sessions, however, so the key is more likely to lie there.

The three strategies most often associated with good generalization (i.e., general case, multiple exemplars, and programming common stimuli) all focus on the role of relevant and irrelevant stimuli within the training situation. Perhaps it is the control of those stimuli which proves most

difficult in the natural environment. For example, while major irrelevant stimuli in a classroom are likely to be associated with the acquisition and practice of many different skills (thereby making them less likely to control any given performance), training a person to purchase food in a single store might consistently pair many irrelevant stimuli with that task and make generalization to other stores less likely. That hypothesis is lent some credence by the fact that the strategy associated with the fourth highest frequencies of generalization (i.e., sequential modification across different situations) would tend to reduce that problem in the natural environment and, indeed, was the only situation in which at least nonseverely handicapped subjects trained in the natural environment generalized better than those trained in an artificial environment.

Finally, in at least some cases, "natural events" may not actually occur or be allowed to occur in the natural environment. One is not likely to allow a trainee to run into traffic during a street-crossing program just to ensure "introduction to the natural contingencies." Also, to improve the frequency of practice, one might be tempted to simply go back and forth across the same street, thereby obviating the natural positive consequence of that activity (i.e., getting closer to one's destination). Such difficulties might account for the fact that there were no substantial differences in the success of "introducing subjects to natural contingencies" within natural and artificial environments.

The provision of training in the natural environment is clearly not a simple proposition. As mentioned at the beginning of this article, such training does help to ensure that trainees will acquire a skill of immediate functionality in at least one situation. Training in most natural situations also affords an increased opportunity for severely handicapped and nonhandicapped persons to interact. If a simple train-and-hope approach is adopted in the natural environment, however, it may result in a substantial decrease in the probability of skill generalization beyond the training situation. Those problems can apparently be avoided through careful planning and the adoption of one or more strategies specifically designed to facilitate generalization.

References

Brown, L., Nisbet, N., Ford, A., Sweet, M., Shiraga, B., and Loomis, R. (1982). *The critical need for nonschool instruction in educational programs for severely handicapped students.* Unpublished manuscript, University of Wisconsin-Madison and Madison Metropolitan School District (U.S. Office of Special Education, Contract No. G008102099).

Hawkins, R. P., and Dotson, V. A. (1975). Reliability scores that delude: An Alice in Wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. In E. Ramp and G. Semb (Eds.), *Behavior Analysis: Areas of research and application.* Englewood Cliffs, NJ: Prentice-Hall.

Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behavior in severely handicapped youth: Pinball machine use. *Journal of the Association for Persons with Severe Handicaps, 6,* 38-44.

Lancioni, G. E. (1982). Normal children as tutors to teach social responses to withdrawn mentally retarded schoolmates: Training, maintenance and generalization. *Journal of Applied Behavior Analysis, 15*(1), 17-40.

Neef, N. A., Iwata, B. A., & Page, T. J. (1978). Public transportation training: In vivo versus classroom instruction. *Journal of Applied Behavior Analysis, 11*, 331-344.

Reese, S. C., & Filipczak, J. (1980). Assessment of skill generalization: Measurement across setting, behavior, and time in and educational setting. *Behavior Modification, 4*, 209-223.

Sailor, W., & Guess, D. (1983). *Severely Handicapped Students: An Instructional Design*. Boston: Houghton-Mifflin Co.

Sprague, J. R., and Horner, R. H. (1984). The effects of single instance, multiple instance, and general case training on generalized vending machine use by moderately and severely handicapped students. *Journal of Applied Behavior Analysis, 17*(2), 273-278.

Stokes, T. F., and Baer, D. B. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

White, O. R. (1984). Selected issues in program evaluation: Arguments for the individual. In B. K. Keogh (Ed.), *Advances in special education, Vol. 4*. Greenwich, CT: JAI Press Inc.

SIX

RESPONSE COMPETITION AND GENERALIZATION

Barbara Matlock
Felix F. Billingsley
and Marsha Thompson

It is currently well-known that various forms of inappropriate behavior (e.g., stereotypic movements and aggression) can be decreased by the development of other behaviors which are considered more desirable. Procedures such as DRI (differential reinforcement of incompatible behaviors) and DRA (differential reinforcement of alternate behaviors) have frequently been used to eliminate inappropriate behavior while increasing the frequency of appropriate behavior (e.g., Favell, 1973; Tarpley & Schroeder, 1979). Those procedures are based on the assumption that the more desirable behavior provides an option to, and/or is structurally (i.e., topographically) incompatible with, the undesirable behavior.

The development of desirable behavior, however, does not assure the permanent elimination of undesirable behavior. When reinforcement for the desired behavior is withdrawn, or when the opportunity to perform the desired behavior is not the most efficient way to access reinforcers, the undesired behavior may, in effect, successfully "compete" with the desired behavior and reappear. Favell, McGimsey, and Schell (1982) reinforced toy play as an alternative to self-injurious behaviors as a means of gaining sensory stimulation in the case of six profoundly retarded adolescents and young adults. A reduction in self-injurious behaviors occurred when toys were available but not when the toys were removed. O'Brien, Azrin, and Bugle (1972) implemented a training program for four profoundly retarded children to increase the ease and speed of walking relative to crawling. As a result of the program, all four children began to walk. Walking, however, appeared to be a less efficient means of gaining reinforcement for two of the children who reverted to crawling following discontinuation of training. Restraint at the waist for 5 seconds following 3 seconds of crawling was used effectively to reduce the efficiency of crawling and promote walking for those two subjects.

The studies cited above suggest that response competition may interfere with the maintenance of newly trained responses. It is also possible that such competition may influence the outcome of attempts to promote generalized effects across behaviors and/or situations and should receive consideration in the development of instructional programs for persons with severe handicaps. This paper presents several studies from the late 1960s to the early 80s which are illustrative of the role of response competition within generalization contexts.

Treatment Generalization Across Behaviors

In a study by Eason, White, and Newsom (1982), six children who had been diagnosed as autistic, or retarded with autistic features, were trained to play with toys appropriately in a playroom setting. The children were then observed with those same toys in an observation room without an adult, in a classroom, and in follow-up playroom situations. It was found that the appropriate toy play behavior increased while stereotypic behavior decreased in all settings even though structural incompatibility did not necessarily exist between the toy play and the stereotypes. Maintenance was also observed across time. The authors speculated "that increased performance of toy play may have introduced children to external and internal 'natural maintaining contingencies of reinforcement'" (Stokes & Baer, 1977, p. 166), or that toy play provided sensory reinforcement that was "better" in quantity and/or quality than that provided by stereotypic behavior. They cautioned that characteristics of the children involved (i.e., all displayed some toy play skills prior to the study) may have also influenced the generalized treatment effect across behaviors. It should be noted that, while the experimental design was sufficient to demonstrate the functional relationship of the treatment to the observed negative covariation in the target behaviors, it was inadequate to demonstrate a functional relationship between the treatment and cross-situational performance (i.e., generalization across settings).

A physical restraint plus reinforcement procedure was used by Whitman, Hurley, Johnson, and Christian (1978) with a severely retarded boy to increase instruction-following behavior and to decrease noncompliance and inappropriate play responses. Three other inappropriate behaviors (i.e., aggression, clothes stripping, and annoying vocalizations) were monitored but not treated.

The child was given activities by his mother during half-hour experimental sessions. When noncompliance or inappropriate play occurred, the mother restrained the child by holding him in a chair for 5 seconds. Positive reinforcement (e.g., juice and praise) was given when the child followed his mother's verbal directions.

The rate of instruction-following and on-task behaviors increased during treatment periods and targeted undesirable behaviors were reduced to a near zero level. Reduction in the untreated aggression and clothes stripping behaviors was also noted. While the amount of annoying vocalizations did not decrease, the topography did alter, from screaming and whining to "gibberish."

One possible explanation mentioned by the authors for the observed generalized treatment effects across behaviors was that sufficient response exemplars were trained (i.e., after having two behaviors consistently punished the child learned that related behaviors would not be tolerated). A second possibility was that aggression and stripping followed noncompliance in a behavioral chain, and that successfully treating one behavior caused the chain to be broken. In relation to the second explanation, it seems likely that the untreated behaviors served an attention-getting function. If that was the case, the new appropriate behaviors could have come to serve that same

function in an equally or more reliable and efficient manner. The behavioral chain, then, could have been broken largely as a result of response competition. In addition to generalization effects across behaviors, informal observation and subjective reports indicated situational generalization in the subject's school and ward environments.

Carr and Kologinsky (1983) developed a procedure that reliably facilitated the use of spontaneous sign language for three children who displayed behaviors characteristic of autism and exhibited poor communication skills. In addition, Carr and Kologinsky were interested in the relationship between spontaneous use of signs and stereotypic behaviors.

In the first of two experiments, each child was taught 10 signs. The children were then required to use the signs to gain access to reinforcement. For two children displaying high rates of stereotypic behaviors, observational data on those behaviors were collected in addition to data related to sign use.

The authors found an inverse relationship between signing and stereotypic behaviors. As spontaneous sign use increased, stereotypic behaviors decreased. The authors, therefore, noted a generalized treatment effect.

Two factors were outlined as possible causes for the response generalization produced by the treatment: (1) functional (rather than topographical) incompatibility between classes of behaviors and (2) reinforcer consistency. Carr and Kologinsky suggested that sign use and stereotypic behaviors competed for reinforcement and that sign use obtained the more potent reinforcer. It was also possible that sign use produced more consistent access to reinforcers than did the stereotypic behaviors.

Generalization Across Situations

Horner (1971) employed rearrangement of the natural environment as a strategy for promoting generalization with a 5-year-old moderately retarded boy with spina bifida. The purpose of training was to teach the child, who either scooted in a sitting position or pulled himself along in a prone position, to walk with the aid of crutches. A 10-step successive approximation sequence was employed to teach him to first walk with the help of parallel bars and then to walk with the crutches. When the training was completed and the child was walking with crutches at criterion level, Horner implemented a contingency management program within everyday living situations. That condition was designed to promote generalization by allowing the child access to meals, play area, school, speech therapy, and a bus ride only if he walked to each activity (i.e., with crutches). This method, in which reinforcers naturally occurring in the environment were made available for walking, appeared extremely successful in establishing cross-situational generalization.

Another study that assessed the effect of manipulation of the nontraining environment on generalization and maintenance was conducted by Stolz and Wolf (1969). A 16-year-old moderately retarded male, who had been described as organically blind, served as the subject. Following visual dis-

crimination training, two naturally occurring situations were observed in order to determine whether environmental manipulation would force the child to respond to new visual cues. In one situation the child was required to obtain his own food items unassisted. In the second situation he was trained to eat appropriately by the provision of instruction in appropriate eating, and criticism and brief periods of food withdrawal for inappropriate eating. In both situations it was found that the subject began to use additional visual cues, but only after changes were made in the nontraining environment. Unfortunately neither Horner (1971), nor Stolz and Wolf used designs that would allow for an experimental analysis of their strategy to promote generalization.

An 11-year-old severely retarded boy, who when asked a "difficult" question (i.e., a question that he could not answer correctly) would frequently echo the question, was the subject of an investigation by Tucker, O'Dell, and Suib (1978). A reinforcement procedure was employed to increase the use of an "I don't know" response when the subject was presented with a difficult question and the echolalia was punished with a loud verbal "NO" and restatement of the question. The training was conducted in a daily experimental session. Results during the session indicated an increase in appropriate responding accompanied by a decrease in the echolalia.

Generalization of the appropriate "I don't know" response within the subject's regular daycare setting was not observed until the staff either prompted the "I don't know" response or punished the echolalia using the procedure employed in training. The reinforcement procedure was deemed unnecessary in the nonexperimental setting. Although this study seemed to indicate that echolalia successfully competed with appropriate responses within the regular daycare setting until the subject was prevented from escaping difficult questions by using echolalia, the research design employed did not permit a true experimental test of that hypothesis.

Using an 11-year-old autistic boy as their subject, Horner and Budd (1983) examined the comparative effects of training setting on sign use, and the relationship between an adaptive behavior (i.e., sign use) and the reduction of potentially competing behaviors (i.e., grabbing and yelling). The child was taught five signs as appropriate responses to questions presented by the trainer, first in a simulated setting and then in the natural classroom setting. Within both simulated and natural settings, grabbing and yelling behaviors were regarded as errors and corrected through restatement of the question and physical prompting of the appropriate sign. If the correct sign was produced, the child was rewarded with praise and the target item (e.g., juice) which was the object of the trainer's question.

The results of training in the simulated setting had virtually no effect on either the use of signs, or on grabbing and yelling, in the classroom. Training in the natural setting, however, was found to be "functionally related both to the use of signs across the school day and a dramatic reduction in the grabbing and yelling" (p. 2). It was noted by the authors that natural setting training may have promoted generalized performance by both strengthening the control of naturally occurring stimuli over sign use, and diminishing the strong control exerted by classroom stimuli over grabbing and yelling.

A Classic Failure to Obtain Generalization

In a study by Risley (1968) several methods were used to eliminate the disruptive and dangerous climbing behavior of a brain damaged, hyperactive 6-year-old girl. Timeout procedures were implemented in the home while extinction and DRI procedure were used in the laboratory. The method used at home involved a 10-minute timeout after every occurrence of climbing. The mother was also instructed to limit physical and verbal interaction as she led the child to the timeout setting and to increase interactions when the child was not climbing. In the laboratory, the opportunity to climb was removed and the child was reinforced for sitting in her chair and looking at the experimenter, an assumed prerequisite to learning imitative behaviors.

Once sitting in the chair and looking at the experimenter (i.e., "behaviors incompatible with climbing," p. 25) had been established, the child was again given the opportunity to climb. At that point, climbing immediately reappeared. The experimenter ignored the climbing behavior and reinforced the child only when she was seated and looking at him. Climbing continued at a high rate. Meanwhile, at home, the timeout procedure was unsuccessful in reducing the climbing behavior.

As Risley noted, "It did not appear that the climbing behavior was maintained by consequences which the experimenter could manipulate" (p. 25). At home, for example, the author suggested that climbing gained parental attention and interaction, which he surmised was reinforcing, but when reinforcement was (presumably) reduced through timeout, no decrease in climbing occurred. In the laboratory, an attempt to provide the child with a more appropriate behavior that would supplant climbing may have failed because the behavior chosen did not compete for the same reinforcers. Climbing in the laboratory was finally eliminated by the contingent application of electric shock. Suppression of climbing was not achieved at home, however, until shock was applied within that setting.

Conclusions

The studies reviewed in this paper were selected to highlight a variable which we feel has been frequently neglected in efforts to promote generalized effects across behaviors and the performance of skills across situations. That variable involves the competition for reinforcers which may exist between newly trained desirable behaviors and existing undesirable behaviors.

Generalization Across Behaviors

Typically, guidance regarding the application of DRI and DRA procedures has emphasized the structural (i.e., topographic) incompatibility of behaviors, or the overall increased density of reinforcement subjects receive for appropriate behaviors (e.g., see Alberto and Troutman, 1982; Dietz and Repp, 1983; and Sulzer-Azaroff and Mayer, 1977). Although differential

reinforcement strategies have undeniably been successful in producing increases in appropriate behaviors and concurrent decreases in inappropriate behaviors in many cases, a number of dramatic failures have been described in the literature (e.g., Risley, 1968). It may be that generalization of treatment effects could be enhanced if incompatible behaviors were selected on the basis of the function they serve, rather than on the basis of (or in addition to) other considerations such as structural incompatibility (cf. Donnellan, Mirenda, Masaros, & Fassbender, 1984). Where Behavior A (i.e., a behavior to be trained as an alternative response) secures the same reinforcing effect as Behavior B (i.e., an existing inappropriate behavior), it seems reasonable to hypothesize that the development of Behavior A would more likely be accompanied by reliable decrements in Behavior B than if it did not allow the subject to achieve the same effect. In addition, treatment generalization across behaviors might be particularly pronounced where Behavior A permits more efficient or reliable access to the same class of reinforcing events as Behavior B. It is possible that the successful competition of appropriate with inappropriate behaviors for such reinforcers was the basis for the generalized treatment effects obtained by workers such as Carr and Kologinsky (1983); Eason, White, and Newsom (1982); and Whitman, Hurley, Johnson, and Christian (1978) cited in this review. In any case, we are in agreement with the opinion of Voeltz, Evans, Derer, and Hanashiro (1984) that "the most lasting approach to decreasing a behavior is the provision of positive alternatives which accomplish the same function" (p. 22), and feel that methods for both the accurate identification of behavioral functions and the selection of the most effective incompatible/alternative behaviors are deserving of future research. Recent guidelines provided by Donnellan et al. (1984) represent a significant step toward the development of a technology for decreasing inappropriate behaviors based on functional incompatibility.

Cross-Situational Skill Performance

Although exhibiting various degrees of experimental rigor, studies by Horner (1971), Horner and Budd (1983), Stolz and Wolf (1969), and Tucker, O'Dell, and Suib (1978) suggest that existing undesirable behaviors which are under the control of stimuli in nontraining settings may interfere with the cross-situational generalization of newly trained behaviors which achieve similar reinforcing effects. While Horner and Budd employed retraining in the natural setting as a means of producing cross-situational responding, the results of the other studies cited indicate that it may be sufficient to reduce the relative efficiency or reliability with which the competing undesirable behavior achieves reinforcing effects in order to produce generalization of the trained behavior. The sufficiency of such a tactic is supported by findings of research sponsored by UWRO (Billingsley & Neel, 1985; Neel & Billingsley, 1984). Where cross-situational generalization fails to occur, then, it may be that the problem (at least in some cases) relates to a failure or inability to deny access to reinforcers for undesirable behaviors in nontraining settings rather than to an instructional deficiency in the training setting. In such cases, generalization might be enhanced by reducing the functionality of the undesirable behavior by ensuring that it does not allow the pupil to achieve reinforcing effects or, as suggested by White (in press), by increasing the efficiency or reliability with which the

trained behavior permits access to reinforcing effects (e.g., the pupil might be trained to perform the behavior more fluently). From a competing behaviors perspective, Risley's (1968) inability to achieve cross-situational suppression of climbing without resorting to electric shock in each setting is not surprising, given that the function served by climbing was never accurately identified and behaviors were not trained which would successfully compete with climbing, even in a structured laboratory situation.

References

Alberto, P. A., & Troutman, A. C. (1982). *Applied behavior analysis for teachers: Influencing student performance*. Columbus, OH: Charles E. Merrill.

Billingsley, F. F., & Neel, R. S. (1985). *Competing behaviors and their effects on skill generalization and maintenance*. Manuscript submitted for publication.

Carr, E. G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis, 16*, 297-314.

Deitz, D. E. D., & Repp, A. C. (1983). Reducing behavior through reinforcement. *Exceptional Education Quarterly, 3*(4), 34-46.

Donnellan, A. M., Mirenda, P. L., Mesaros, R. A., & Fassbender, L. L. (1984). Analyzing the communicative functions of aberrant behavior. *Journal of the Association for Persons with Severe Handicaps, 9*, 201-212.

Eason, L. J., White, M., & Newsom, C. (1982). Generalized reduction of self-stimulatory behavior: An effect of teaching appropriate play to autistic children. *Analysis and Intervention in Developmental Disabilities, 2*, 157-169.

Favell, J. (1973). Reduction of stereotypes by reinforcement of toy play. *Mental Retardation, 11*, 21-23.

Favell, J. E., McGimsey, J. F., & Schell, R. M. (1982). Treatment of self-injury by providing alternate sensory activities. *Analysis and Intervention in Developmental Disabilities, 2*, 83-104.

Horner, R. D. (1971). Establishing the use of crutches by a mentally retarded spina bifida child. *Journal of Applied Behavior Analysis, 4*, 183-189.

Horner, R. H., & Budd, C. M. (in press). Teaching manual sign language to a nonverbal student: Generalization and collateral reduction of maladaptive behavior. *Education and Training of the Mentally Retarded*.

Neel, R. S., & Billingsley, F. F. (1984, November). *Form, function, and the selection of alternative behaviors*. Paper presented at the Eighth Annual Conference on Severe Behavior Disorders of Children and Youth, Tempe, AZ.

O'Brien, F., Azrin, N. H., & Bugle, C. (1972). Training profoundly retarded children to stop crawling. *Journal of Applied Behavior Analysis, 5*, 131-137.

Risley, T. R. (1968). The effects and side effects of punishing the autistic behaviors of a deviant child. *Journal of Applied Behavior Analysis, 1*, 21-34.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Stolz, S. B., & Wolf, M. M. (1969). Visually discriminated behavior in a "blind" adolescent retardate. *Journal of Applied Behavior Analysis, 2*, 65-77.

Sulzer-Azaroff, B., & Mayer, G. R. (1977). *Applying behavior analysis procedures with children and youth*. New York: Holt, Rinehart, and Winston.

Tarpley, H. D., & Schroeder, S. R. (1979). Comparison of DRO and DRI on rate of suppression of self-injurious behavior. *American Journal of Mental Deficiency, 84*, 188-194.

Tucker, G. H., O'Dell, S. L., & Suib, M. (1978). Control of selective echolalia via the instatement of a general alternative response. *Behavior Research and Therapy, 16*, 302-306.

Voeltz, L. M., Evans, I. M., Derer, K. R., & Hanashiro, R. (1984). Targeting excess behavior for change: A clinical decision model for selecting priority goals in educational contexts. *Child and Family Behavior Therapy, 5*(3), 17-35.

White, O. R. (in press). Aim*star wars (setting aims that compete). Episode II: Return of the learner. *Journal of Precision Teaching*.

Whitman, T. L., Hurley, J. D., Johnson, M. R., & Christian, J. G. (1978). Direct and generalized reduction of inappropriate behavior in a severely retarded child through a parent-administered behavior modification program. *AAESPH Review, 3*, 68-77.

SEVEN

TEACHING RETARDED STUDENTS TO REINFORCE THEIR OWN BEHAVIOR: A REVIEW OF PROCESS AND OPERATION IN THE CURRENT LITERATURE

Kathleen A. Liberty and Lawrence J. Michael

There has been increasing discussion of the possibilities and potential benefits of teaching retarded students to use self-control techniques such as self-reinforcement (Bernstein, 1981; Holman & Baer, 1979; Hops, 1983; Kurtz & Neisworth, 1976; Mickler, 1984; Thomas, 1980). Over the last two decades researchers have shown that self-reinforcement can increase the fluency with which a previously acquired behavior is performed (Ballard & Glynn, 1975; Rosenbaum & Drabman, 1979; Kazdin, 1978; O'Leary & Dubey, 1979), is useful in maintaining performance levels achieved during more traditional treatments and interventions (Bolstad & Johnson, 1972; Meichenbaum & Goodman, 1971; Rosenbaum & Drabman, 1979; Turkewitz, O'Leary, & Ironsmith, 1975), and may mediate contingency differences to promote cross-setting generalization of performance (Bornstein & Quevillon, 1976; Rosenbaum & Drabman, 1979). As an added advantage, self-reinforcement techniques may require less teacher time than other types of interventions (Bolstad & Johnson, 1972; Turkewitz, O'Leary, & Ironsmith, 1975). In short, advocates of self-reinforcement strategies with retarded citizens can cite much experimental support.

Catania (1975) has identified two crucial dimensions to self-reinforcement: the *process* of self-reinforcement, in which the question is whether the act of self-reinforcement produces an effect on the target behavior, and the *operation* of self-reinforcement, in which the behaviors that constitute the act of self-reinforcement are examined. A review of the current literature was undertaken to examine the experimental evidence that has accumulated concerning the effects on the behavior of retarded subjects of the process of self-reinforcement, and to identify how subjects reinforce their own behavior and the training procedures involved in the acquisition of the operation of self-reinforcement.

Method

Articles selected for the first stage of the review were those in which the subjects were identified as mentally retarded and in which typical applied behavior analysis methodology was employed (i.e., articles which did not present repeated measures per phase were not analyzed, including: Helland, Paluck, & Klein, 1976; Robertson, Simon, Pachman, & Drabman, 1979; and

Shapiro & Klein, 1980). Nine studies involving 7 mildly retarded, 9 moderately retarded, 3 severely retarded subjects, and 1 profoundly retarded subject were identified. An additional 14 subjects were identified as either falling in the mildly or moderately retarded classifications (Frederiksen & Frederiksen, 1975).

The methodology sections of these studies were examined to identify descriptions of those factors critical to an analysis of the process of self-reinforcement: (a) the *target response* (i.e., the response which was to be changed by the process of self-reinforcement) and (b) the *controlled response* (i.e., the response which was reinforced by the subject). Several analysts have identified methodological factors which might confound the reported effects of the process of reinforcement (Catania, 1975; Gross & Wojnilower, 1984; Jones, Nelson, & Kazdin, 1977; Kazdin, 1978; Kazdin, 1980). The presence or absence of such factors may not account for the reported results, but may provide cautionary information in interpreting results. The presence or absence of three major factors was evaluated in each of the studies: (a) whether a change was made in the type or contingency of reinforcement at the same time as self-reinforcement was introduced, (b) whether additional potentially reinforcing events were introduced concurrent with self-reinforcement, and (c) whether self-reinforcement was included in a "package" of self-control techniques.

The review of the operation of self-reinforcement and its training was extended to include three articles involving retarded subjects that were eliminated from the first analysis (i.e., Helland, Paluck, & Klein, 1976; Robertson, Simon, Pachman, & Drabman, 1979; and Shapiro & Klein, 1980). In order to determine any differences between retarded and nonretarded subjects in the operation of self-reinforcement, 11 studies involving nonretarded subjects and a variety of operational responses were also added to the review. The methodology sections were examined to identify (a) the schedule of self-reinforcement, (b) the discriminative stimuli (S^D) for the operation of self-reinforcement (i.e., what signaled to the subject to begin the operation of self-reinforcement), (c) a precise description of the *controlling response* (i.e., the behaviors that constituted the operation of self-reinforcement), and (d) the type of reinforcer that was delivered. Since many of the articles involved the delivery of secondary reinforcers, both immediate and delayed controlling responses and reinforcers were identified.

Several analysts have identified the components sufficient and necessary to discriminate the operation of self-reinforcement from other forms of behavior (Bandura, 1976; Catania, 1975; Kazdin, 1978). The operation of self-reinforcement was reviewed to determine (a) whether or not the subjects adopted performance standards that determined the criteria for reinforcement; (b) whether or not the subjects had full control over the reinforcers, which were freely available during the entire experimental period in which the subject was performing the controlled response; and (c) whether or not subjects actually administered the reinforcers contingent upon performance of the controlled response (i.e., if data included in the study showed the accuracy or reliability of self-reinforcement delivery) and, conversely, withheld reinforcers for noncriterion behavior.

Studies were examined to identify the antecedent and consequent events used in training, the amount of time spent in training, and the criteria used for concluding training. In addition, we noted whether or not data on the independent (i.e., unprompted) operation of self-reinforcement or on the reliability of self-reinforcement were included, since these measures would permit an analysis of the effectiveness of the training procedures.

Results

The Process of Self-Reinforcement

The results of this analysis are shown in Table 7-1. Two studies reported accuracy on tasks. Accuracy increased for 3 subjects (Burgio, Whitman, & Johnson, 1980; Hanel & Martin, 1980) and either did not change or decreased for 7 subjects (Hanel & Martin, 1980).

All of the nine articles included at least one target behavior for which the aim was either to increase the fluency of a desirable behavior or to decrease the fluency of an undesirable behavior. In all but two of these studies, the fluency of the target behavior was effected in the desired direction; in Burgio, Whitman, and Johnson (1980) the rate of academic performance generally decreased for 2 subjects during self-reinforcement (contingencies were for accuracy, not rate, however), and in Bates, Renzaglia, and Clees (1980), self-reinforcement alone had no effect on the rate of producing drapery pulleys (fluency did improve, however, once changing criteria were introduced concurrent with self-reinforcement).

Two studies reported maintenance of effects for at least one target behavior. A reduced rate of verbal disruptions by workers maintained for six months (Gardner, Cole, Berry, & Nowinski, 1983) and for one year (Gardner, Clees, & Cole, 1983).

Gardner, Clees, and Cole (1983) reported that self-reinforcement for a response class which included not engaging in high-rate disruptive vocalizations resulted in a concurrent deceleration in stereotypic head and/or hand and arm shaking or flapping. Burgio, Whitman, and Johnson (1980) reported generalized improvement in performance to an untrained task and in on-task behavior to an untrained classroom setting.

Methodological confounds limit the generality of effects in all of the reviewed studies with retarded subjects. The effects of self-reinforcement were confounded by its integration in a self-control package in six of the nine studies. For example, Jackson and Martin (in press) taught subjects to set goals and monitor their behavior, as well as to deliver tokens to themselves. In seven studies, the self-reinforcement condition was contrasted with other experimental phases in which reinforcement was delivered by an external agent (e.g., the experimenter or teacher). In six of these, the effects of the process of self-reinforcement were confounded by changes or additions in reinforcement that accompanied the introduction of the phase which included self-reinforcement. For example, Hanel and Martin (1980) taught

Table 7-1
The Process of Self-Reinforcement

Article	Subjects	N[1]	Target Response	Controlled Response	Reported Effect on Target Response[2]	Confounding Factors
Bates, Renzaglia, & Clees (1980)	Profound	1	Vocational assembly task	Production rate	No effect	
Burgio, Whitman, & Johnson (1980)	Mild Moderate	1 1	Off task Task accuracy & fluency	On task Arithmetic, printing	Off task reduced; effects generalized across tasks, but not across all settings. Task accuracy increased while fluency generally decreased. Accuracy on a generalization task increased.	In package; change in reinforcement
Frederiksen & Frederiksen (1975)	Mild & Moderate	14	On task Disruptive behavior	Subject controlled (unknown)	On task increased; disruptive behaviors decreased	
Gardner, Clees, & Cole (1983)	Moderate	1	Verbal disruption Talks to others	"Good adult worker" (vocational assembly task) Production rate	Decreased talks to self & talking to others. Generalized decrease of stereotypic behavior. Effects for talks to self maintained at one year. Production rate increased.	In package; change in reinforcement contingency
Gardner, Cole, Berry, & Nowinski (1983)	Moderate	2	Verbal disruptions	"Good adult worker" (vocational assembly task)	Decreased fluency of disruption; maintained at 6 months.	In package; change in reinforcement contingency
Hanel & Martin (1980)	Mild Moderate Severe	3 3 2	Vocational assembly task	Production rate and accuracy	Fluency increased; accuracy increased for 1, unchanged for 5, decreased for 2.	In package, change in reinforcement contingency
Horner & Brigham (1979)	Mild	2	Disruptive behaviors	On task	Fluency decreased	In package; additional reinforcement
Horner, Lahren, Schwartz, O'Neill, & Hunter (1979)	Severe	1	Vocational assembly task	Production rate	Fluency increased	Additional reinforcement
Jackson & Martin (in press)	Mild Moderate	1 2	Vocational assembly task	Production rate	Fluency increased	In package; change in reinforcement contingencies

[1]Subjects actually trained to self-reinforce (excludes control and other untrained subjects).
[2]Effects during phase(s) with self-reinforcement.

BEST COPY AVAILABLE

subjects to both set goals and deliver tokens for performing to or exceeding their goals. Since the contingency changed with the goal, reported effects may be attributed to the changed contingency rather than to the process of self-reinforcement. In Horner, Lahren, Schwartz, O'Neill, and Hunter (1979), the controlling response of the subject resulted in both a token and a bell during the self-reinforcement phase; the bell was not available when the experimenter operated the apparatus. The potential additional reinforcing effects of the bell are not otherwise explored.

In one of the two studies where self-reinforcement followed a reported "unreinforced" baseline, performance changes may be the result of changes in the quantity or quality of reinforcement, as well as the agent of reinforcement. For example, Horner and Brigham (1979) report that contingent reinforcement for on-task behavior was not available during baseline; effects during the self-reinforcement phase may be due to the change in the quantity of reinforcement rather than to the agent of reinforcement.

The Operation of Self-Reinforcement

The results of this analysis are shown in Table 7-2 (studies involving retarded subjects are presented first). Both retarded and nonretarded subjects were taught a variety of immediate controlling responses, including verbal behaviors (e.g., say praise statements), written responses (e.g., write a " + "), operating an apparatus (e.g., pushing a lever or actuating a wrist counter), and taking coins or tokens. Two studies with retarded subjects (Helland, Paluck, & Klein, 1976; Jackson & Martin, in press) included chained responses (e.g., writing a mark and taking a token) as did one study with nonretarded subjects (Morrow & Presswood, 1984). In our sample of studies, only retarded subjects were taught to take tokens directly (Gardner, Clees, & Cole, 1983; Gardner, Cole, Berry, & Nowinski, 1983; Shapiro & Klein, 1980).

For most of the studies in which a delayed controlling response was applicable, some kind of exchange for back-up reinforcers took place. However, authors generally failed to describe whether back up reinforcers were delivered by the subject or by someone else. In three studies subjects delivered back-up reinforcers to themselves, including free time (Glynn, Thomas, & Shee, 1973; Uhlman & Shook, 1976) and food (Horner & Brigham, 1979).

In a little less than one half of the studies, subjects began to self-reinforce when they finished performing the target response. For example, in Hanel and Martin's study (1980), the subjects' completed assembly of an airline coffee pack was the discriminative stimulus for delivery of a token. A bell on a timer or a prerecorded tone on an audio tape was used to signal self-reinforcement in six studies. An interval schedule was used in each of these studies. The subjects actually set the timer or activated the tape recorder in three of these studies (Gardner, Clees, & Cole, 1983; Gardner, Cole, Berry, & Nowinski, 1983; Morrow & Presswood, 1984). In all, subjects controlled the discriminative stimuli in 11 of the 23 studies. In the other studies, external agents set timers, activated tapes, asked subjects questions, or prompted self-reinforcement in some other fashion to provide the S^D for the operation

Table 7-2
The Operation of Self-Reinforcement

			The Operation of Self-Reinforcement				Properties of Self-Reinforcement		
Article	**Subjects**	**Controlled Response**	**Schedule**	**Discriminative Stimuli**	**Controlling Response†**	**Reinforcer†**	**Adopt Standards?**	**Free Access?**	**Contingent Delivery?**
Bates, Renzaglia, & Clees (1980)	Profound	Vocational assembly task	FR 2	Complete task put in block (block has room for 2). When 2 transfer to box.	**Immediate**: take penny from container	**Immediate**: penny	yes	yes	unknown
			FI 80'	Teacher's comments	**Delayed**: exchange	**Delayed**: lunch snack items	no	no	unknown
Burgio, Whitman, & Johnson (1980)	Mild Moderate	Academic tasks: arithmetic, printing, phonics	Mixed VR VI subject-control during & after	Perform task	Say praise	Praise	yes	yes	no
Frederiksen & Frederiksen (1975)	Mild Moderate	Unknown (subject controlled)	FI 30'	Teacher asks "Did you earn a token?"	**Immediate**: say "yes" [say "no"]	**Immediate**: teacher gives token [no token]	no	no	yes
			Unknown		**Delayed**: exchange	**Delayed**: free time & privileges			
Gardner, Clees, & Cole (1983)	Moderate	Vocational assembly task	FI 2' to FI 60'	Timer rings verbally label behavior as "adult worker" or "not adult "worker"	**Immediate**: take coin from clip on card on desk	**Immediate**: penny or coin? photo of subject smiling	yes	limited	unknown
			Unknown		**Delayed**: spend during break	**Delayed**: not reported	unknkown	unknown	unknown
Gardner, Cole, Berry, & Nowinski (1983)	Moderate	Vocational assembly task	FI 3' to FI 60'	Timer rings verbally label behavior as "adult worker" or "not adult "worker"	**Immediate**: take coin from clip on card on desk	**Immediate**: penny or coin? photo of subject smiling	yes	limited	unknown
			Unknown		**Delayed**: spend during break	**Delayed**: not reported	unknown	unknown	unknown
Hanel & Martin (1980)	Mild Moderate Severe	Vocational assembly task	FR 1	Finish task	**Immediate**: push lever on apparatus	**Immediate**: blue or orange marble	yes	yes	yes
			FI 20' (session)		**Delayed**: exchange marbles	**Delayed**: orange marble = 1 penny	unknown	unknown	unknown
			am FI (2-3 sess) pm FI (4-5 sess)		**Delayed**: spend at workshop "store."	**Delayed**: food at retail cost	unknown	yes?	unknown

†Overt responses and presumed reinforcing events only.

(table continues)

The Operation of Self-Reinforcement (continued)

Article	Subjects	Controlled Response	The Operation of Self-Reinforcement				Properties of Self-Reinforcement		
			Schedule	Discriminative Stimuli	Controlling Response	Reinforcer	Adopt Standards?	Free Access?	Contingent Delivery?
Holland, Paluck, & Klein (1976)	Mild	Vocational assembly task	FR 10	Colored slip of paper inserted in work stack	Say praise & take dime or candy from pile	Praise & dime or candy	yes	yes	no
Horner & Brigham (1979)	Mild	On task	FI 5	Timer rings	**Immediate**: write a "+" or a "-" in a box on a sheet (six boxes/sheet)	**Immediate**: "+"	yes	limited (to 6)	yes
			FI 30'		**Delayed**: go to back of room & get candies from file cabinet	**Delayed**: 1 "+" = 1 candy	no	limited	unknown
Horner, Lahren, Schwartz, O'Neill, & Hunter (1979)	Severe	Vocational assembly task	FR 1	Finish task check quality	**Immediate**: push lever on apparatus	**Immediate**: bell & token	yes	yes	yes
			FI 90'		**Delayed**: exchange	**Delayed**: food	unknown	unknown	unknown
Jackson & Martin (in press)	Mild Moderate	Vocational assembly task	Changing criterion (set by subject)	Finish task	**Immediate**: check lists on form, take chip from pile, place in cup	**Immediate**: check mark & chip	yes	yes	yes
			End of day		**Delayed**: exchange chips	**Delayed**: quarter	unknown	unknown	unknown
Robertson, Simon, Pachman, & Drabman (1979)	Mild Moderate Severe	Appropriate behavior	FI 10' free per morning	Timer rings	**Immediate**: say number of points (operate apparatus to award points: 2 = "good", 1 = "ok", 0 = "not good")	**Immediate**: points (visible on apparatus) (10 max per day)	yes	unclear	unknown
			FI 30'		**Delayed**: exchange points at store in classroom	**Delayed**: edibles, activities, privileges (variable point cost)	unknown	unknown	unknown
Shapiro & Klein (1980)	Mild	On task Pre-academic skills	Mixed FI 30', FI 60', FI 90'	Timer rings; teacher asks "Are you working?" Prompts to take token (faded).	**Immediate**: take token from box in front of subject	**Immediate**: token	yes	yes	unknown
			FI 30' (if earn 10 of 15 available tokens)		**Delayed**: exchange	**Delayed**: subjects choice of objects	no	unknown	unknown

(Overt responses and presumed reinforcing events only.

(table continues)

BEST COPY AVAILABLE

The Process of Self-Reinforcement (continued)

Article	Subjects	Controlled Response	The Operation of Self-Reinforcement: Schedule	Discriminative Stimuli	Controlling Response†	Reinforcer†	Properties of Self-Reinforcement: Adopt Standards?	Free Access?	Contingent Delivery?
Santogrossi & Clement (1985)	Non-hand	Attend to math	VI ?	Tone on audio tape	**Immediate**: push button on bar	**Immediate**: point at flash of light from box	no	yes	yes
			FI 10'		**Delayed**: exchange points	**Delayed**: 1 point = 1 piece of food	unknown	unknown	unknown
Santogrossi, Bartell, & McLeod (1979)	Non-hand	Attending	Changing subject—set criterion for on task Unknown	Apparatus in subject's pocket signals intervals	**Immediate**: actuate wrist counter **Delayed**: exchange	**Immediate**: number on counter = points **Delayed**: toys, special trips	yes unknown	yes unknown	unknown unknown
Glynn, Thomas, & Shee (1973)	Non-hand	On task	VI ? Unknown	Tone on audio tape	**Immediate**: write a check in a square **Delayed**: subject goes to recess early	**Immediate**: 1 checkmark **Delayed**: 1 check = 1 minute	yes yes	yes limited (to recess)	yes unknown
Harris (in press)	Learning Disabled Non-hand	Puzzle assembly	Mixed VR VI subject control during & after responding	Perform task	Say 'praise'	'Praise'	yes	yes	unknown
Harris & Graham (in press)	Learning Disabled	Writes words in story	Mixed VR VI subject control during & after responding	Perform task	Say 'praise'	'Praise'	yes	yes	unknown
Mischenbaum & Goodman (1971)	Behavior probs.	Attention to task	Mixed VR VI subject control during & after responding	On task	Say 'praise'	'Praise'	unknown	yes	unknown
Morrison & Pressnall (1984)	Multi-hand (Severe beh. dis. & deaf)	Not engaging in stereotypic behaviors	Not reported (VI ?)	Timer rings & happy face light on experimenter-controlled apparatus (Unhappy face light)	Sign 'praise' & mark '+' on a grid (Sign 'feedback' & mark a '−')	'Praise' & '+'	no	yes	no

†Short responses and presumed reinforcing events only.

(table continues)

BEST COPY AVAILABLE

The Operation of Self-Reinforcement (continued)

Article	Subjects	Controlled Response	The Operation of Self-Reinforcement: % Include	Discriminative Stimuli	Controlling Response?	Reinforcer?	Properties of Self-Reinforcement: Adapt Standards?	Free Access	Contingent Delivery?
Rhode, Morgan, & Young (1983)	Behav. probs.	Appropriate behavior	FI 15' to FI 60' then VR 2 days	Teacher signal	**Immediate**: mark card up to 5 points (& say points—faded)	**Immediate**: points	yes	yes	yes
			End of session; fade to VR 2 days		**Delayed**: exchange (faded)	**Delayed**: small toys & food (faded)	unknown	unknown	unknown
Turkewitz, O'Leary, & Ironsmith (1975)	Academic & Social probs.	Reading Appropriate behavior	FI 15	Teacher signal	**Immediate**: give (?) up to 5 pts for work, up to 5 pts for behavior	**Immediate**: points	no	yes	unknown
			FI 60		**Delayed**: exchange at class "store" (faded)	**Delayed**: toys & food (1 pt = [illegible] [item = retail cost ÷ 1.3] (faded)	unknown	limited	unknown
Ulman & Mank (1979)	Multi-hand	Task completion	FR 1	Finish task to turn task card free time card on desk)	**Immediate**: unknown	**Immediate**: unknown	unknown	unknown	unknown
			Unknown		**Delayed**: take free time	**Delayed**: 10' free time with activities & tasks	unknown	yes	unknown
Lahey & Drake (1979)	Hyper-active	Academic tasks	VI	Finish work period	**Immediate**: look at clock on desk, actuate wrist counter once for each point to be given for that period	**Immediate**: number on counter = number of points	yes	yes	unknown
			Subject set criterion	Reach criterion	**Delayed**: exchange	**Delayed**: subject selected objects & activities	yes	unknown	unknown

BEST COPY AVAILABLE

of self-reinforcement. For example, in Frederiksen and Frederiksen (1975), the teacher asked the subjects whether they had earned a token at the end of a 30-minute work period.

There was little difference in the nature of the reinforcers delivered by retarded and nonretarded subjects. Praise alone reinforced subjects in 4 studies, while retarded subjects in 1 study immediately received a tangible unconditioned primary reinforcer (Helland, Paluck, & Klein, 1976). Subjects in 16 of the 22 studies received a conditioned reinforcer immediately; usually marks or tokens for retarded subjects, and points for nonretarded subjects (immediate reinforcer not identified by Uhlman and Shook, 1976).

When specified, delayed reinforcers provided to retarded subjects were similar to those provided nonretarded subjects. Food was the most frequent delayed reinforcer (five studies), followed by free time (three studies), and money (two studies). Combinations of these, including also toys and the subjects' choice of activities were used in six studies while in two studies delayed reinforcers were not specified or were not described.

The operation of immediate self-reinforcement occurred on a fixed interval schedule in eight studies and intervals ranged from 30 seconds (Shapiro & Klein, 1980) to 60 minutes (Gardner, Clees, & Cole, 1983; Gardner, Cole, Berry, & Nowinski, 1983). Subjects in five studies used a fixed ratio and subjects in three studies applied a variable interval schedule. In two studies, schedules changed with criterion (Fantuzzo, Harrell, & McLeod, 1979; Jackson & Martin, in press). Mixed variable ratio/variable interval schedules were used only when the controlling response was praise (four studies). The schedule of self-reinforcement was not reported in one study.

Only 11 of the 17 studies with delayed reinforcement included its schedule. Of those 11, 7 reported that an exchange of tokens for back-up reinforcers took place at the end of the session (i.e., fixed interval). To fade back-up reinforcers, Rhode, Morgan, and Young (1983) gradually changed the schedule of delayed reinforcement from the end of every session to a variable ratio 2 days. In Hanel and Martin's study (1980), subjects exchanged reinforcers on two separate occasions following the initial immediate self-delivery of marbles (FR1). At the end of each session subjects exchanged marbles for pennies (FI20'), and at the end of each morning and afternoon work period they exchanged pennies for food (FI 2 to 4 sessions).

An analysis of the properties of immediate self-reinforcing operations showed that, in all but one study (Frederiksen & Frederiksen, 1975), subjects had free, if sometimes limited, access to the immediate reinforcers. In all but four studies (Fantuzzo & Clement, 1981; Frederiksen & Frederiksen, 1975; Morrow & Presswood, 1984; Turkewitz, O'Leary, & Ironsmith, 1975), subjects adopted performance standards which defined when the controlling response should occur. However, only eight studies reported data on whether or not subjects correctly and contingently self-reinforced. In one of the eight studies (Glynn, Thomas, & Shee, 1973), accuracy data were collected for only 5 days because observers found it difficult to code concurrently with other data. Reported accuracy ranged from 76% for 8 nonretarded subjects (Glynn, Thomas, & Shee, 1973) to 100% for 1 nonretarded subject (Fantuzzo & Clem-

ent, 1981). The accuracy of retarded subjects ranged from 80% (Horner & Brigham, 1979) to 99% (Horner, Lahren, Schwartz, O'Neill, & Hunter, 1979; Jackson & Martin, in press). Six of the 23 studies were judged as meeting all three of the criteria necessary for defining the property of immediate self-reinforcement. For example, Hanel and Martin (1980) taught subjects to push a lever when they completed assembly of an airline coffee pack (i.e., adopt standards), and, although subjects were free to push the lever and deliver a marble (i.e., token) to themselves at any time (i.e., free access), they reinforced themselves contingently between 88% and 98% of the time (although they sometimes required retraining). In contrast, retarded subjects in Frederiksen and Frederiksen (1975) were not trained to differentiate on-task from off-task behavior (i.e., no standards), they had access to tokens only at periods determined by the teacher (i.e., no free access), and a token was delivered if the student said he had earned one. The authors report that students "almost always" said "yes." And, in fact, student on-task behavior increased. However, it is impossible to determine from the study whether students said "yes" following an on-task period and said "no" following an off-task interval. Catania (1975) would be unlikely to find any face validity in this operation of self-reinforcement. The operation of the delayed reinforcers in studies was usually insufficiently explained for analysis of their properties.

Training Self-Reinforcement

Methods of training subjects to self-reinforce were rarely described so that specific antecedents and consequences could be identified or procedures could be replicable. It appears, however, that generally similar training conditions were applied with retarded and nonretarded subjects. In studies with mentally retarded subjects, training antecedents consisted of unspecified "verbal instructions" (Bates, Renzaglia, & Clees, 1980; Hanel & Martin, 1980; Horner & Brigham, 1979; Robertson, Simon, Pachman, & Drabman, 1979) or of "demonstrations" with verbal instructions (Gardner, Clees, & Cole, 1983; Gardner, Cole, Berry, & Nowinski, 1983; Helland, Paluck, & Klein, 1976; Jackson & Martin, in press; Shapiro & Klein, 1980). Consequences during training included verbal feedback (Gardner, Clees, & Cole, 1983; Hanel & Martin, 1980; Robertson, Simon, Pachman, & Drabman, 1979), presumably for failing to self-reinforce, and praise for appropriate self-reinforcement (Horner, Lahren, et al., 1979). Horner, Lahren, et al. also dismantled the 10-part test adapter just assembled by the subject if the subject failed to deliver reinforcement contingently. In addition, Robertson, et al. (1979) gave an M & M for the students' accurate matching of reinforcement with the teacher and took away a point for inaccurate matching.

Burgio, Whitman, and Johnson (1980) described training procedures of a different nature and in more detail. First, while the subject observed, the experimenter performed the controlled response while verbalizing the controlling responses in a self-control package. Next, the subject performed the controlled response while the experimenter verbalized the controlling responses. And finally, the subject performed the controlled response and verbalized the controlling responses as the experimenter whispered along with the subject. Training included both instances of the behavior to be

reinforced and the behavior which was not to be reinforced. Subjects continued to receive training until they performed the entire sequence of controlling responses accurately and independently for both correct and incorrect controlled responses during three consecutive training sessions. Then subjects were trained to perform under various distracting conditions.

Only four other articles described the criteria used for terminating training. Three and five consecutive correct trials on all self-control behaviors were required by Gardner, Cole, Berry, and Nowinski (1983) and Gardner, Clees, and Cole (1983), respectively. Horner and Brigham (1979) required five consecutive correct trials at each of increasingly longer intervals (i.e., from one to five minutes) before concluding training. While Jackson and Martin (in press) required four consecutive correct trials of one self-control component before moving on to the next training step, no criterion was specified for terminating training of the last self-control step.

Five articles with retarded subjects reported the amount of training time, which seemed to increase with the reported classification of retardation. When subjects were trained to implement a self-control package which included self-reinforcement, mildly retarded subjects spent averages of 1.3 hours (Horner & Brigham, 1979) and 2.0 hours in training (Helland, Paluck, & Klein, 1976) while moderately retarded subjects were trained for 4.0 hours (Gardner, Cole, Berry, & Nowinski, 1980) and for 5.5 hours (Gardner, Clees, & Cole, 1983). Hanel and Martin (1980) reported an average of 2.7 hours for a mixed group of mildly, moderately, or severely retarded subjects. A severely retarded subject was taught to self-reinforce in 9.0 hours (Horner, Lahren, Schwartz, O'Neill, & Hunter, 1979). In contrast, a behaviorally disordered youth was taught to operate a self-reinforcement apparatus in 10 trials over a 0.5 hour period (Morrow & Presswood, 1984), and Meichenbaum and Goodman (1971) taught behaviorally disordered and/or low IQ (above 85) second-graders a self-control package in 2.0 hours. Data illustrating the acquisition of the controlling responses during training were not included in any of the articles reviewed.

Discussion

Given the interest in self-reinforcement procedures, it was surprising that we were able to identify only nine studies in the recent literature in which data showed the effects of the process of self-reinforcement over time on the target behaviors of mentally retarded individuals. However, the effects of self-reinforcement with these 34 subjects were very similar to results with other groups of subjects (Meador & Ollendick, 1984; O'Leary & Dubey, 1979; Rosenbaum & Drabman, 1979). In general, teaching retarded individuals to reinforce their own performance may be of questionable value during the acquisition of skills in which accuracy of performance is critical. This is not unexpected, since in such a situation the student must simultaneously acquire the discriminative stimuli and contingencies for a minimum of two chained responses (i.e., the controlled and the controlling responses).

Self-reinforcement seems to be most useful in accelerating or decelerating previously acquired target behaviors in which the primary concern is the frequency of performance. This is in keeping with general instructional research with severely handicapped students, which suggests that changes in reinforcers are most likely to improve performance when the response is fluent, while more precise discriminative stimuli will generally improve nonfluent and inaccurate performance (Haring, Liberty, & White, 1980). These results seemed to hold true for all levels of retardation; however, the small number of subjects represented by the reviewed articles certainly limits this generalization.

While the appeal of self-reinforcement for promoting maintenance and generalization is great, very little experimental data can be cited to support it, as is the case with other groups of subjects (Meador & Ollendick, 1984; O'Leary & Dubey, 1979). On the other hand, no evidence contraindicating self-reinforcement with retarded subjects was found, nor do the data suggest that the behavior of individual retarded subjects was effected any differently than that of anyone else. If desirable and even socially important performance changes may result, then the procedures provide an alternative to external control tactics that may be useful to practitioners.

A very limited range of target behaviors for mentally retarded subjects was included in the reviewed studies. On-task, disruptive behaviors, and vocational assembly tasks were the only target behaviors studied. While the choice of target behaviors seems to parallel those represented in studies with other groups of students (Meador & Ollendick, 1984; O'Leary & Dubey, 1979; Rosenbaum & Drabman, 1979), research with critical response classes, including communication, locomotion, and social interaction skills, would broaden the empirical basis for the application of self-reinforcement with mentally retarded individuals.

An evaluation of the efficacy of self-reinforcement with retarded subjects must also be tempered by the methodological confounds identified in the studies reviewed. These problems are evident in the experimental literature in self-control, and are not just limited to studies with retarded subjects (Gross & Wojnilower, 1984; Kazdin, 1978; Kazdin, 1980; Jones, Nelson, & Kazdin, 1977; Meador & Ollendick, 1984). A study which is confounded by additional external reinforcement introduced at the same time as self-reinforcement, or one in which self-reinforcement is one factor in a self-control package, does not permit a functional analysis of the process of self-reinforcement. This is critical from an experimental point of view, but may not be as important from a practical perspective. If one cannot identify which component of a package is responsible for performance changes, then the practitioner must teach all components; if future research identifies those components both sufficient and necessary for performance change, then instruction can be modified to include only effective components.

Retarded subjects acquired a variety of operations for self-reinforcement, and, according to the few articles that supplied data, were able to reinforce their own behavior as accurately as did other groups of subjects. Unfortunately, most articles failed to report these data. Reporting procedural reliability documents the contingent use of self-reinforcement procedures

(Billingsley, White, & Munson, 1980) and helps define the operation of self-reinforcement (Bandura, 1976; Catania, 1975). High levels of reliability also help substantiate a meaningful relationship between the process of self-reinforcement and changes in the controlled response.

In a recent review, Gross and Wojnilower (1984) suggest that subjects who are actually free to self-reinforce without environmental constraints for accurate and contingent delivery may reward themselves noncontingently (i.e., "cheat"). They state that an externally controlled contingency for accurate self-reinforcement "must be applied" (Gross & Wojnilower, 1984, p. 509). Robertson et al. (1979), Turkewitz et al. (1975), and Horner, Lahren, et al. (1979) reported using some contingencies for noncontingent self-reinforcement. If such contingencies must be maintained indefinitely, then the advantages of self-reinforcement are substantially reduced, since an external agent must always be present to enforce the contingencies and, by extension, self-control is not practiced. For example, Robertson et al. (1979) used such a contingency but they were able to eventually fade the external controls on "cheating" while maintaining accurate and "fair" self-delivery. Most studies, however, did not describe how external controls were faded.

Instruction in the standards for contingent delivery of self-reinforcement is part of the acquisition of the controlling response (Catania, 1975). The failure of students to apply standards accurately is a failure of the instructional procedures used to teach the operation of self-reinforcement. Most authors failed to explicitly describe how subjects were taught to reinforce themselves contingently or to document the acquisition of independent, unprompted, and reliable controlling responses. Thus it is impossible to determine the differences between those methods which apparently produced highly accurate self-reinforcement (e.g., Jackson & Martin, in press) and those which produced much less accurate delivery (e.g., Glynn, Thomas, & Shee, 1973). In addition, the lack of detail in the methodological sections of most of the reviewed articles prevents a useful and/or replicable description of how to teach the operation of self-reinforcement. Studies that did provide some indication as to instructional method generally described procedures which are fairly typical, including the use of external agents to deliver teaching antecedents and consequences. A few studies did mention that the externally controlled antecedents and consequences were faded (e.g., Hanel & Martin, 1980), or that instruction continued until subjects performed independently (e.g., Horner & Brigham, 1979). More precise published descriptions of methodology, and the inclusion of data on the acquisition of the controlling response and on its performance once contingencies for it were withdrawn, would permit a clearer understanding of the operation of self-reinforcement. Such descriptions would also provide information needed to determine methods of instructing students to reinforce their own behavior.

Two additional issues with the operation of self-reinforcement suggest fundamental problems in instructing the controlling response. First, several authors report that subjects were retrained in the controlling response during the course of the study (Hanel & Martin, 1980; Morrow & Presswood, 1984; Stevenson & Fantuzzo, 1984; Turkewitz, O'Leary, & Ironsmith, 1975), indicating the failure of the teaching methods to result in maintenance of that response. Second, in some studies, external constraints on the accuracy

of self-reinforcement were never faded (e.g., Horner & Brigham, 1979, retained control of back-up reinforcers throughout their study) and, in these studies, independent and accurate self-reinforcement was never achieved.

While one may assume that the operation of self-reinforcement may be instructed as any other skill, one of the major reasons for the interest in self-reinforcement and other types of self-control is the failure of methods which rely on external control of antecedents and consequences to produce maintained and generalized skills (Gross & Wojnilower, 1984; Meador & Ollendick, 1984). It seems likely that the application of traditional "train and hope" methodologies, which rely completely on externally controlled antecedents and consequences, would subvert the development of maintained and generalizable self-controlling responses just as they do for most other behaviors (Stokes & Baer, 1977; White, Leber, & Phifer, Chapter 5 in this publication). If the major aim of teaching self-control is to teach the student to be the agent of change of his or her own behavior, then the controlling response itself must maintain and generalize beyond the training setting and after training ceases. Therefore, we must identify methods of instruction which develop not only accurate and fluent controlling responses, but maintained and generalized self-control skills.

One procedure which has produced generalized controlling responses has been identified by Meichenbaum (Meichenbaum, 1977; Meichenbaum, 1979; Meichenbaum & Goodman, 1971). Burgio, Whitman, and Johnson (1980) used this procedure to teach mildly or moderately retarded subjects a series of controlling responses. They provide data on the percentage of intervals during which the subjects were observed to perform the controlling responses in two different nontraining settings. In the first transfer setting, both subjects showed rapid acceleration and consistency in performance of the controlling responses. In the second transfer setting, controlling responses occurred much less frequently and with greater variability, perhaps, as the authors suggest, because of the "inhibitory" effect of the classroom setting on audible controlling responses (Burgio, Whitman, & Johnson, 1980). Meichenbaum (1977) has suggested a final step in this kind of training, in which the subject whispers to himself, and then performs the sequence of controlling responses covertly. Burgio, Whitman, and Johnson (1980) did not include this final step, since they observed and reported whether the subjects engaged in the controlling responses (i.e., covert responses can not be measured by outside observers). The experimenters also attempted to measure the maintenance of the controlling responses once training was faded; however, the study ended before training concluded for subjects. Unfortunately, the authors did not report the accuracy with which subjects applied standards and reinforced their work either during training or in the generalization settings, nor did they provide data on the acquisition of the controlling responses during training. Despite these limitations, this article indicates an alternative instructional method, which fades the externally controlled instructional antecedents and consequences for the controlling response as it is acquired, and which suggests that generalized, if not maintained, controlling responses are produced. The extension and/or adaptation of such a procedure to other classes of controlled responses, and with other retarded subjects, should surely be investigated.

Although data are scanty for empirically justifying the use of self-reinforcement with retarded persons, its appeal, like that of other self-control behaviors, is unlikely to diminish given the disappointing record of externally controlled interventions to produce maintained and generalized skills by retarded individuals and the difficulty of arranging environmental contingencies outside of training settings. Research in self-reinforcement will contribute information on the efficacy of an alternative intervention. However, unless more precise and fully documented descriptions of actual self-reinforcement operations are included, the information will not be useful in replications or applications. The documentation of training methods and the acquisition of the controlling responses is also critical to the development of tactics both effective and practical in producing accurate, maintained, and generalized self-control skills.

References

Ballard, K., & Glynn, T. (1975). Behavioral self-management in story writing with elementary school children. *Journal of Applied Behavior Analysis, 8*, 387-398.

Bandura, A. (1976). Self-reinforcement: Theoretical and methodological considerations. *Behaviorism, 4*, 135-155.

Bates, P., Renzaglia, A., & Clees, T. (1980). Improving the work performance of severely/profoundly retarded young adults: The use of a changing criterion procedure design. *Education and Training of the Mentally Retarded, 15*(2), 95-104.

Bernstein, G. (1981). Research issues in training interpersonal skills for the mentally retarded. *Education and Training of the Mentally Retarded, 16*, 70-74.

Billingsley, F., White, O., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment, 2*, 229-241.

Bolstad, O., & Johnson, S. (1972). Self-regulation in the modification of disruptive classroom behavior. *Journal of Applied Behavior Analysis, 5*, 443-454.

Bornstein, P., & Quevillon, R. (1976). The effects of a self-instructional package on overactive preschool boys. *Journal of Applied Behavior Analysis, 9*, 179-188.

Burgio, L., Whitman, T., & Johnson, M. (1980). A self-instructional package for increasing attending behavior in educable mentally retarded children. *Journal of Applied Behavior Analysis, 13*, 443-459.

Catania, A. (1975). The myth of self-reinforcement. *Behaviorism, 3*, 192-199.

Fantuzzo, J., & Clement, P. (1981). Generalization of the effects of teacher- and self-administered token reinforcers to nontreated students. *Journal of Applied Behavior Analysis, 14*, 435-447.

Fantuzzo, J., Harrell, K., & McLeod, M. (1979). Across-subject generalization of attending behavior as a function of self-regulation training. *Child Behavior Therapy, 1* (4), 313-321.

Frederiksen, L., & Frederiksen, C. (1975). Teacher-determined and self-determined token reinforcement in a special education classroom. *Behavior Therapy, 6*, 310-314.

Gardner, W., Clees, T., & Cole, C. (1983). Self-management of disruptive verbal ruminations by a mentally retarded adult. *Applied Research in Mental Retardation, 4*(1), 41-58.

Gardner, W., Cole, C., Berry, D., & Nowinski, J. (1983). Reduction of disruptive behaviors in mentally retarded adults: A self-management approach. *Behavior Modification, 7*, 76-96.

Glynn, E., Thomas, J., & Shee, S. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis, 6*, 105-113.

Gross, A., & Wojnilower, D. (1984). Self-directed behavior change in children: Is it self-directed? *Behavior Therapy, 15*, 501-514.

Hanel, F., & Martin, G. (1980). Self-monitoring, self-administration of token reinforcement, and goal-setting to improve work rates with retarded clients. *International Journal of Rehabilitation Research, 3*(4), 505-517.

Haring, N., Liberty, K., & White, O. (1980). Rules for data-based strategy decisions in instructional programs. In W. Sailor, B. Wilcox, and L. Brown (Eds.), *Methods of instruction for severely handicapped students*. Baltimore, MD: Paul H. Brookes.

Helland, C., Paluck, R., & Klein, M. (1976). A comparison of self- and external reinforcement with the trainable mentally retarded. *Mental Retardation, 14*(5), 22-23.

Holman, J., & Baer, D. (1979). Facilitating generalization of on-task behavior through self-monitoring of academic tasks. *Journal of Autism and Developmental Disorders, 9*(4), 429-445.

Hops, H. (1983). Children's social competence and skill: Current research practices and future directions. *Behavior Therapy, 14*, 3-18.

Horner, R., & Brigham, T. (1979). The effects of self-management procedures on the study behavior of two retarded children. *Education and Training of the Mentally Retarded, 14*(1), 18-24.

Horner, R., Lahren, B., Schwartz, T., O'Neill, C., & Hunter, J. (1979). Dealing with the low production rates of severely retarded workers. *AAESPH Review, 4*(2), 202-212.

Jackson, D., & Martin, G. (in press). Additive effects of components of a self-control package for improving the production of mentally handicapped workers. *Journal of Practical Approaches to Developmental Handicap*.

Jones, R., Nelson, R., & Kazdin, A. (1977). The role of external variables in self-reinforcement. *Behavior Modification, 2*(2), 147-178.

Kazdin, A. (1978). *History of behavior modification*. Baltimore: University Park Press.

Kazdin, A. (1980). *Behavior modification in applied settings*. Homewood, IL: The Dorsey Press.

Kurtz, P., & Neisworth, J. (1976). Self-control possibilities for exceptional children. *Exceptional Children, 42*, 212-217.

Meador, A., & Ollendick, T. (1984). Cognitive behavior therapy with children: An evaluation of its efficacy and clinical utility. *Child and Family Behavior Therapy, 6*(3), 25-44.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York: Plenum Press.

Meichenbaum, D. (1979). Teaching children self-control. In B. Lahey and A. Kazdin (Eds.), *Advances in clinical child psychology* (Vol 2). New York: Plenum Press.

Meichenbaum, D., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77*(2), 115-126.

Mickler, M. (1984). Self-management skill training for educable mentally retarded. *Journal of Special Education, 18*(2), 143-149.

Morrow, L., & Presswood, S. (1984). The effects of a self-control technique on eliminating three stereotypic behaviors in a multiply-handicapped institutionalized adolescent. *Behavioral Disorders, 9*(2), 131-140.

O'Leary, S., & Dubey, D. (1979). Applications of self-control procedures by children: A review. *Journal of Applied Behavior Analysis, 12*, 449-465.

Rhode, G., Morgan, D., & Young, K. (1983). Generalization and maintenance of treatment gains of behaviorally handicapped students from resource rooms to regular classrooms using self-evaluation procedures. *Journal of Applied Behavior Analysis, 16*, 171-188.

Robertson, S., Simon, S., Pachman, J., & Drabman, R. (1979). Self-control and generalization procedures in a classroom of disruptive retarded children. *Child Behavior Therapy, 1*(4), 347-362.

Rosenbaum, M., & Drabman, R. (1979). Self-control training in the classroom: A review and critique. *Journal of Applied Behavior Analysis, 12*, 467-485.

Shapiro, E., & Klein, R. (1980). Self-management of classroom behavior with retarded/disturbed children. *Behavior Modification, 4*(1), 83-94.

Stevenson, H., & Fantuzzo, J. (1984). Application of the "generalization map" to a self-control intervention with school aged children. *Journal of Applied Behavior Analysis, 17*, 273-278.

Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Thomas, M. (1980). A conversation with Herbert Goldstein about mildly retarded learners. *Education and Training of the Mentally Retarded, 15*, 216-219.

Thoresen, C., & Mahoney, M. (1974). *Behavioral self-control*. New York: Holt, Rinehart and Winston.

Turkewitz, H., O'Leary, K., & Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. *Journal of Consulting and Clinical Psychology, 43*(4), 577-583.

Uhlman, W., & Shook, G. (1976). A method of maintaining high rates of performance in an open classroom setting. In T. Brigham, R. Hawkins, J. Scott, & T. McLaughlin (Eds.), *Behavior analysis in education: Self-control and reading*. Dubuque, IA: Kendall Hunt.

Varni, J., & Henker, B. (1979). A self-regulation approach to the treatment of three hyperactive boys. *Child Behavior Therapy, 1*(2), 171-192.

White, O., Leber, D., & Phifer, C. (in press). Training in the natural environment and skill generalization: It doesn't always come naturally. In N. Haring (Principal Investigator), *Investigating the problem of skill generalization* (3rd ed.). (US Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

EIGHT

SOCIAL SKILLS TRAINING REVIEW

Greg Weisenstein
Sharon Field
and Carol Kiolet

Many authors have noted the deficits in social skills possessed by severely handicapped persons, as well as the significant negative consequences of those deficits (e.g., Berler, Gross, & Drabman, 1982; Cone, Anderson, Harris, Goff, & Fox, 1978; Senatore, Matson, & Kazdin, 1982; Vaughn, Ridley, & Cox, 1983). Fortunately, "during the past several decades, it has been demonstrated that severely handicapped students can be taught to display a variety of positive social behaviors. . . . However, there have been conflicting results in terms of the newly acquired social behaviors generalizing outside of the direct intervention setting" (Stainback, Stainback, & Strathe, 1983, p. 293). "An important consideration . . . is whether the trained social behavior was maintained outside of the initial training sessions. Thus far, the findings for generalization have been less promising than those for initial training. When generalization had been achieved by some training procedures, the frequency of social behavior was marginal and was maintained only by the continuation of reinforcement" (Peterson, Austin, & Lang, 1979, p. 82). "Obviously, generalization is critical, since social behaviors that are not exhibited across different people in a variety of settings are of limited value. For example, 'cooperative' social behavior is of limited utility when a child learns to display such behavior with his handicapped classmates in a special education class but does not display the behavior with other people in other settings" (Stainback, Stainback, & Strathe, 1983, p. 293).

Voeltz (1984) discusses the importance of social skills training for integrating persons who are severely handicapped into the community and states that "traditionally, social skills have been a neglected area in curricula for severely handicapped children and youth" (p. 166). The author describes a social performance curriculum model utilized by the Hawaii Integration Project which includes three major features:

> 1) a skill acquisition perspective, as opposed to the incorporation of a deviance reduction component; 2) a view of social skills that involves not only exhibiting appropriate behavior (or withholding certain behavior) but also making important discriminations among multiple cues that indicate which behaviors are appropriate; 3) the interrelatedness of social skills with responses from other domains. (p. 167)

The model also divides all social skill task performance contexts into eight major interactive situation types: (1) Private Independence: Free Time and Task Related; (2) Friendship Interaction: Free Time and Task Related; (3) Homogeneous Small Group: Play, Social, Task Related; (4) Heterogeneous

Small Group: Play, Social, Task Related; (5) Active, Goal-Oriented Group; (6) Passive, Goal-Oriented Group; (7) Public Place Stranger Interactions, Social and Task Related; and (8) Public Independence: Free Time and Task Related (p. 169). Within this model each social situation is analyzed according to major environmental, situational, relationship, and internal cues provided by the specific situation. Through a Special Friends program, peer "interactions are designed as 'generalization sessions' that provide an additional, more natural opportunity for the handicapped child to practice, and the teacher to measure, the acquisition of skills as outlined in the student's IEP" (p. 170).

The variation of cues which are situation specific is recognized in this model and the implications for generalization are discussed. Programming for generalization of social skills is also provided in the model. However, no data are provided as to the success of these generalization efforts.

Gaylord-Ross and Pitts-Conway (1984) also emphasize the importance of the development of social behavior for autistic persons. According to Gaylord-Ross and Pitts-Conway, "Social behavior is the instructional domain that will be most critical in determining whether autistic individuals succeed in less restrictive, more normalized settings" (p. 198). A model for integration of autistic students at the high school level is described. They state that the major thrust of previous research in autism has been centered on the process by which autistic persons learn and perform. "Content, like social responses or fine motor responses, has served as a means to examine the way that antecedent and consequent events influence the expression of behavior" (p. 199).

Gaylord-Ross and Pitts-Conway delineate four main contexts for social interactions on a continuum according to the degree of structure inherent in the situation. The most highly structured type of social interaction is peer tutoring followed by leisure exchanges and transient interactions. Student-centered interactions are the most loosely structured type of social interaction. The authors make the point that more structured social situations often form the basis for more informal relationships and interactions.

Programming for generalization is described in this model. "In the Marin County program, we first work on generalizing the response to multiple people and settings through use of a simultaneous training procedure. Here, the student is taken to at least three settings to practice the greeting response. In each setting, at least three nondisabled persons are approached with a greeting. During generalization training a time delay . . . procedure could be used to progressively fade out the trainer" (p. 211).

They state that it is also important to "teach discrimination of the transient response. The practical meaning of discrimination in the case of greetings is that the student does not offer greetings to every passer-by. For instance, it may be proper to greet a friend at the beginning of lunch, but it would be inappropriate to repeatedly greet the person every time you pass him during lunch. More importantly, the context in which one meets strangers determines whether a greeting is delivered. For example, in a crowded city street or in a public bathroom it would be unwise to train autistic adolescents to

make greeting responses. In contrast, at a party or in a small, familiar store it could be appropriate to initiate greetings. The number of contexts where it might be appropriate or inappropriate to initiate greetings are myriad. Further demonstration and research work is needed to delineate social discrimination patterns among this group of students" (p. 212).

Gaylord-Ross and Pitts-Conway emphasize the importance of social skills acquisition and describe some strategies for generalization of social skills. However, no data is provided as to the success of these generalization efforts.

Timm, Strain, and Eller (1979) conducted research to increase the levels of positive social behavior. Their subjects were three socially withdrawn preschool boys. The design for this study was a combination of withdrawal of treatment and multiple baseline procedures. During the first intervention, subjects received a fixed number of prompts and contingent attention events for positive social behavior (which was not operationally defined in the published article). Prompts and contingent attention events were then reduced on a response-dependent basis for two subjects, and reduced on a response-independent basis for the third subject.

> The results suggest that: (a) the intervention procedures produced marked increase in positive social behavior emitted by each subject; (b) response-dependent fading and thinning, contrasted with response-independent tactics, maintained levels of positive social behavior equivalent to those observed; . . . (c) changes in positive and negative behaviors emitted by peers paralleled changes in positive and negative behaviors emitted by each subject; and (d) no 'spillover' of treatment effects was noted for subjects during periods in which they were not direct recipients of intervention procedures. (p. 308)

While we see that response-dependent fading and thinning was effective in maintaining positive social behavior in spite of reduced prompts and reinforcers, this study is a prime example of the criticisms of Peterson, Austin, and Lang (1979) as there was no generalization to "no reinforcement" conditions.

Also supporting this criticism is a study by Cone, Anderson, Harris, Goff, and Fox (1978). This study focused "on . . . increasing associative and cooperative play, and on documenting correlated changes in self-stimulation, aggression, and inactivity" (p. 352). The subjects for this study were five profoundly retarded males who were residents of the same living unit of a state facility for retarded children. They were selected based on aide reports that each rarely interacted socially. The average chronological age of the subjects was 15.3 years, and all five had intellectual functioning within the profound range. The design for this study was a combined multiple baseline (with subjects being grouped as two pairs and a single) and withdrawal design, involving the following phases:

Baseline. Observers merely scored the behavior of the five children in the large play area of their living unit.

Ball-toss training off unit. Each child was verbally prompted to throw a ball to the trainer, and then physical and verbal prompts were used to get the children to throw the ball to each other. This training was conducted in an area outside of the living unit.

Generalization I. "Because ball tossing and interaction began occurring at high rates in the training sessions but not in the play area of the living unit, an effort was made to transfer stimulus control to that area. . . . The ball-toss trainer was also present in the play area, though he interacted only minimally with the five children and never prompted or rewarded ball tossing" (p. 354).

Rewarded generalization I. "The trainer began verbally prompting ball tossing for the first two boys and rewarded it with M & M's during the observations in the play area" (p. 354). An important note about the procedure here is that off-unit training sessions were discontinued at this point for the first subjects and simultaneously begun with the next subjects, who also received ball-toss prompting and rewarding in the living unit play area.

Generalization II. This was identical to the Generalization I phase.

Rewarded generalization II. Both the first subjects and the second subjects were again verbally prompted and rewarded for ball-tossing, while the last subject was introduced to these procedures for the first time, and without any prior off-ward training.

Reduced rewards. "During both rewarded generalization phases, ball tosses were followed by an M & M to the tosser on a continuous reinforcement (CRF) schedule. . . . [During this phase] the schedule was changed to approximately a variable ratio two (VR2) for the first four boys and was gradually reduced to a single contingent presentation of an M & M. . . . The single contingent M & M was presented in a different, randomly selected 15-second interval each session" (p. 354-355).

Follow-up.

> All five children were placed in other social interaction training groups, which employed the procedures of the rewarded-generalization and reduced-rewards conditions to increase general social interaction. The five boys were periodically placed back in the play area by themselves to assess maintenance of the changes produced in the earlier phases. Postprogram checks occurred 10, 23, 26, and 37 days after formal sessions had ended. During these four follow-up sessions, a single response-contingent M & M was again presented during a randomly selected 15-second interval. (p. 355)
>
> [The results indicate that] social interaction, defined as associative and cooperative play, was increased in these profoundly retarded institutionalized young males, thereby replicating and extending the work of earlier investigators. . . . Unlike the results of the previous studies, however, consistent increases in social interaction in the living area were not automatically associated with training occurring somewhere else. It was necessary to prompt and reinforce ball tossing

> in the living area before systematic increases were observed. Indeed, the increase in social interaction in the fifth boy without formal training indicates such training may not have been necessary at all. . . . It may be sufficient merely to reinforce and fade the reinforcers in the setting in which increased interaction is desired. The importance of developing an effective social-behavior-shaping technology for institutionalized retarded children is clearly supported by the present data showing correlated changes in other, less adaptive responses. (pp. 357-359)

Vaughn, Ridley, and Cox (1983) dealt with subjects with relatively high intellectual functioning. "Interpersonal problem-solving" training programs have been used successfully to teach conflict resolution and effective interaction skills with a variety of special populations. "The purpose of this study was to evaluate the effects of an interpersonal problem-solving skills program [i.e., the Vaughn and Ridley program] on the interpersonal skills of mentally retarded students" (p. 191).

This study utilized a group design with 30 subjects. The subjects were from a self-contained special school in Australia for students who are mentally retarded. The Behavioral Interpersonal Problem Solving Test (BIPS) was used as the measure of interpersonal skills. The BIPS was administered to the experimental group before intervention, as well as to a control group. The testing situation was as follows: One of the subject's peers was given a toy and told not to give it to the subject. The subject was told that the peer had had the toy for a long time and asked for possible ways to gain access to the toy. Each subject was presented with between seven and ten problem situations, each having the same format with the exception of the desired object which changed with each situation.

Experimental subjects then participated in the interpersonal problem-solving skills training, while control subjects participated in reading story sessions to control for contact. Posttests were then given to experimental and control subjects. "Results indicated a significant treatment effect at posttest on relevant solutions . . . [and] that the experimental group, relative to the contact control group, demonstrated a significant increase in interpersonal problem-solving behavior" (p. 194).

We see again that social skills training proved effective, but there was no examination of generalization effects of any sort presented in this study, which is a prime example of the criticisms of Stainback, Stainback, and Strathe (1983) regarding the relevance of the targeted behaviors for subjects in their daily lives.

Berler, Gross, and Drabman (1982) used the skills training approach to improve the interactions of socially unskilled children on several variables:

> including role-play performance, frequency of verbal and play interactions in a naturalistic free play setting, and peer acceptance as determined by sociometric ratings. Rather than follow a "train and hope"

> model of generalization, the treatment program . . . incorporated several procedures for the purpose of promoting generalization to the natural environment. (p. 42)

The subjects were six boys attending a school for learning disabled children in Jackson, Mississippi, who were selected on the basis of teacher referrals, as having poor peer relationships, and sociometric ratings (i.e., classmates had ranked them last as desirable students to work and play with).

> The two children from each of the three classes who received the lowest mean rating from all classmates based on the six administrations of the questionnaire and were identified by the teacher as having poor peer relationships were selected to participate in the study. One child from each class was randomly assigned to the experimental group and the second child from the class was then assigned to the control group. (p. 43)

Skills training for each experimental subject, which was conducted in a role-play setting, was evaluated with a multiple baseline across two categories of behavior, eye contact and appropriate verbal content in several areas of responding. Duration of speech served as an untrained corollary measure. The components were trained in a sequential and cumulative fashion in a randomly determined order. A role-play test was developed to assess the skills training program.

The investigators attempted to program generalization in their training procedures through:

(a) using two trainers to increase the diversity of stimulus conditions during training;
(b) using a teacher as a second trainer to provide a common stimulus across the training setting and the daily school environment;
(c) group training to allow the children's peers to serve as stimuli common to the training setting and the natural environment;
(d) devoting the last part of each training session to rehearsing responses to scenes spontaneously developed by the children to try to make the role-play scenes relevant to their own interpersonal experiences; and,
(e) verbal and written instructions given to teachers asking them to provide daily feedback to subjects (both experimental and control) concerning their general interactions with peers, paying particular attention to the target behavior being taught in the group sessions at that time.

Experimental and control subjects were assessed on the following measures of generalization:

1. Role-play posttests on the second and fourth day following the termination of treatment, consisting of the 12 trained scenes and 8 untrained scenes.

2. "The social interactions of each child were observed in an in vivo free play situation during baseline and the week following the termination of treatment. Observations were made during the children's daily, unstructured recess period" (p. 46).

3. Sociometric questionnaires, which were readministered to the three classes on four occasions during the 2-week period following the termination of treatment.

4. "In order to assess the maintenance of treatment effects, the behavior of the experimental and control children was assessed on all dependent measures beginning approximately 1 month following the termination of treatment" (p. 46).

> A moderate degree of generalization was demonstrated on untrained role-play scenes as well as when testing was conducted by novel experimenters who approximated the age of the subjects. In addition, the subjects' performance was generally maintained at 1 month follow-up.
>
> . . . the present research also attempted to assess generalization and validate the improvement shown on the role-play tests by measuring the children's social competence on a number of criterion variables, including overt behavior in a naturalistic free play setting and peer sociometric ratings. Improved social behavior in the school environment was not supported on the basis of these generalization measures. Thus, in comparison to three socially unskilled learning disabled children who did not receive training, the experimental subjects in the present study showed no treatment-related changes in their social status or frequency of verbal and play interactions with peers. (pp. 49, 51)
>
> [The authors point out] the selection of target behaviors based on performance in the natural environment may be more relevant to a child's daily social interactions than would behaviors based on performance deficits in role-play situations. Furthermore, target behaviors that have an empirically demonstrated relationship with a criterion measure of social competence (such as sociometric ratings) may be the most valid and relevant skills to be focused on during training. . . . Social skills training programs that do not show generalization to the natural environment should be looked upon critically and should not be considered to be effective until proven otherwise. (p. 52)

Further implications for the social skills training setting were found by Foxx, McMorrow, and Schloss (1983). The researchers modified the board game "Sorry" such that the cards used contained a social situation designed to elicit a complex verbal response from the player. For example, "Someone tells you that your drawing isn't very good . . . What would you do?" The six social skills target areas were compliments, social interaction, politeness, criticism, social confrontation, and questions/answers. The design for this study was a multiple baseline across two groups (three subjects per group).

The results indicated that the intervention improved social skills in all six target areas. In addition, the results generalized first to a casual interview with a novel person in a new environment, then to a novel person in a new environment in a novel situation (i.e., a confederate acted out a situation on a game card, with the subject being led to believe the encounter was spon-

taneous). Furthermore, casual observations, reductions in medication, and progress made by subjects towards community integration (which was dependent on social skills) were offered to "cautiously be considered as indirect evidence for generalization" (p. 166).

The authors offer several interpretations of the data, but regarding generalization in particular point out that the study "conducted the training in a situation (i.e., the game) where social interaction with peers was likely to occur. In other words, the game included many of the characteristics that other studies have built into their analogue settings. . . . Perhaps the failure of many social skills studies to report greater generalization of behaviors is a result of the lack of similarity between the training and natural settings" (p. 169).

Peterson, Austin, & Lang (1979) conducted a study designed to increase the frequency of social behavior exhibited by subjects, as well as examine generalization of effects. "The three participants, selected initially on the basis of teacher recommendation were visually impaired, severely and profoundly retarded adolescents from a special-education class for visually impaired, multihandicapped, low-functioning, nonacademic children (Nashville, Tennessee). The teacher had been concerned because these adolescents infrequently engaged in social interaction with peers and felt that they would possibly benefit from some intervention focusing on socialization" (p. 83). The design of this study was a single-subject, reversal design with replication across subjects for training and generalization involving the following phases:

Pretraining period. During this phase the teacher engaged the target children in their typical daily activities in a small-group setting, making no attempts to foster social activities among the subjects.

Training period. Prompting and teacher praise were used to increase the level of the subject's interaction with peers. "Examples of teacher prompts were: 'Show _____ the letters,' 'Would you like to play ball with _____?' and 'How would you call _____ on the phone?'" (p. 83).

Generalization period. The subjects were integrated with classmates not participating in the study while the teacher was absent.

> Results of the present study show that, given specific intervention procedures, the frequency of social activity found among severely and profoundly retarded adolescents can be increased. Moreover, and most importantly, these individuals can be expected to generalize their newly acquired social skills to a free-play setting in which no reinforcers are given. . . . The subjects' regular teacher acted as the trainer, and all sessions took place in their regular classroom. It is logical to assume that the familiarity of the social context would act as an initial prerequisite for fostering increased social interactions. This phenomenon would have a definite impact not so much during training sessions, but more so during generalization when no prompting or reinforcement is given.

> [Furthermore,] an individual who begins to exhibit more social interactions with peers will have the effect of stimulating those peers to increase their rate of social behaviors toward him or her. As a result, mutually reinforcing, reciprocal behavior is established. The data on the adolescents in this study suggest that such events may, in fact, be occurring. (p. 85)

The authors attribute their positive findings for generalization to factors similar to those found by Foxx, McMorrow and Schloss (1983). These factors are not excluded from other studies which have failed to find evidence for generalization (e.g., Berler, Gross, & Drabman, 1982), thus supporting the contention of Stainback, Stainback, and Strathe (1983) that the results regarding generalization of social skills training are conflicting.

Gaylord-Ross, Haring, Breen, and Pitts-Conway (1984) conducted two experiments to study the initiation and duration of social interactions between autistic and nonhandicapped youths. The first experiment taught two autistic youths to utilize appropriate leisure objects (e.g., a radio, a video game, and gum). The students also received instruction in social skills. A variety of persons (i.e., multiple exemplars) was utilized to teach the skills. Generalization probes were conducted sequentially as training proceeded. For both subjects it was found that in the generalization setting, no social skills initiation occurred during baseline and very few social interactions were initiated during the object only phase. Initiation of social interactions by the autistic individual increased dramatically during social skills training. In addition, nonhandicapped peers approached the autistic youths more frequently after social skills training than they did during the object only phase. The authors attribute the generalization to nontraining contexts to the use of sufficient exemplars during the training setting.

A second experiment was conducted to replicate the effects of the training package with another autistic student. During the second experiment, the object training phase was combined with the social skill training and only one person was utilized to provide the training. Again, generalization of initiation of social interaction by autistic individuals increased after social skill and object function training. Initiation of social contacts by nonhandicapped peers also increased after social skills and object training.

Senatore, Matson, and Kazdin (1982) examined the component parts of the successful social skills training package.

> The emphasis in treatment has shifted toward a more comprehensive treatment approach to alter a variety of behaviors related to social interaction. Typically, treatment incorporates instructions, performance feedback, modeling, role playing and social reinforcement as a social skills training package . . . Several studies have demonstrated the efficacy of a standard social skills training package for modifying interpersonal skills. . . . Additional work is needed to evaluate whether the social skills training package can be improved. Major impetus for investigating variations of the package is the need to ensure generalization of the treatment beyond the usual assessment

> conditions. . . . The purpose of the present investigation was to evaluate the social skills training package to develop skills of mentally retarded adults. (p. 314)

The subjects for this study were borderline to severely retarded with a mean chronological age of 36. All lived in the community, either with their parents or in a supervised living arrangement.

> The dependent measures consisted of role-play performance of social skills, an interview, and performance at a party. Role-play and interview measures were assessed at pretest, posttest, and follow-up. The party was assessed at posttest only. All of these assessments involved direct behavioral observations. . . . A sample [role-play] scene follows: You have just come to the hospital for a group meeting. You missed the meeting last week and you are wondering if other group members will want to talk to you this week. *Role Model Prompt*: Hi. I'm glad you came today. I missed you last week. *Narrator (to subject)*: You say . . .
>
> The scoring system used for all the dependent variables was based on social behaviors that direct care staff considered most important. (p. 316-317)

Thus, while not empirically shown to be relevant to daily activities, an attempt was made to target responses important to the subjects' daily activities. In addition to the role-play scenes, clients were scored based on their performance in a fairly systematic interview dealing with social situations, and at a party where their performance was scored based on responses to questions asked by two undergraduate students.

Three experimental conditions were compared. First was a no treatment control condition; second was a standard social skills package condition; third was a social skills plus active rehearsal condition, where in addition to instructions, modeling, role-playing, and social reinforcement, the subjects "acted out" the scenes by walking through and overtly rehearsing the situations with the use of prompts by the therapist and props such as tables, chairs, etc. "Acting out scenes was initially done with the therapist serving as both the narrator and role model prompt, which was the procedure always included in the previous treatment group. As client skills improved, they not only served as the respondent but took on the roles as role model prompt" (p. 319).

> [The results indicated that] on each measure the social skills training plus active rehearsal group was significantly higher in social skills than the other two conditions; the standard social skills training conditions was significantly more effective than the no-treatment control group. . . . The present investigation demonstrated that the social skills training package with active rehearsal was more effective than the standard method without such rehearsal. . . . Furthermore, treatment effects generalized to a naturalistic setting (party) and were

maintained at a 6-month follow-up assessment. . . . With regard to generalization, treatment effects were noted in behaviors not treated and in behavior in a naturalistic setting (party). (p. 320-323)

The generalization of social skills training is an important area of research as persons who are severely handicapped are integrated into community settings. The importance of social skills for persons who are severely handicapped has been well documented in the literature. Researchers have also noted the importance of the generalization of those skills. For example, if a student learns to perform a greeting response, but can do so only in the classroom, that skill is of limited value to that individual. Furthermore, if she/he is unable to distinguish among the myriad of social cues in the natural environment which affect the appropriateness of the greeting response in a generalized setting, the skill again is of little value to the individual.

The importance of social skills to persons who are severely handicapped and the complexity surrounding the development of instructional strategies for teaching severely handicapped individuals to generalize social skills makes social skills generalization an important line of research. As researchers develop technology to facilitate generalization, it is imperative that this technology be applied to social skills training. Social skills are of the utmost importance in allowing individuals who are severely handicapped to make the transition from school to work and to live successfully in the community, and, at present, strategies to consistently facilitate the generalization of these skills have not been discovered.

References

Berler, E. S., Gross, A. M., & Drabman, R. S. (1982). Social skills training with children: Proceed with caution. *Journal of Applied Behavior Analysis, 15*(1), 41-53.

Cone, J. D., Anderson, J. A., Harris, F. C., Goff, D. K., & Fox, S. R. (1978). Developing and maintaining social interaction in profoundly retarded young males. *Journal of Abnormal Child Psychology, 6*(3), 351-360.

Foxx, R. M., McMorrow, M. J., & Schloss, C. N. (1983). Stacking the deck: Teaching social skills to retarded adults with a modified table game. *Journal of Applied Behavior Analysis, 16*(2), 157-170.

Gaylord-Ross, R. J., Haring, T. G., Breen, C., & Pitts-Conway, V. (1984). The training and generalization of social interaction skills with autistic youth. *Journal of Applied Behavior Analysis, 17*(2), 229-247.

Gaylord-Ross, R. J., & Pitts-Conway, V. (1984). Social behavior development in integrated secondary autistic programs. In Certo, N., Haring, N., & York, R. (Eds.), *Public school integration of severely handicapped students* (pp. 197-219). Baltimore: Paul H. Brookes.

Peterson, G. A., Austin, G. J., & Lang, R. P. (1979). Use of teacher prompts to increase social behavior: Generalization effects with severely and profoundly retarded adolescents. *American Journal of Mental Deficiency, 84*(1), 82-86.

Senatore, V., Matson, J. L., & Kazdin, A. E. (1982). A comparison of behavioral methods to train social skills to mentally retarded adults. *Behavior Therapy, 13*, 313-324.

Stainback, W., Stainback, S., & Strathe, M. (1983). Generalization of positive social behavior by severely handicapped students: A review and analysis of research. *Education and Training of the Mentally Retarded, 18*(4), 293-299.

Timm, M. A., Strain, P. S., & Eller, P. H. (1979). Effects of systematic, response-dependent fading and thinning procedures on the maintenance of child-child interaction. *Journal of Applied Behavior Analysis, 12*(2), 308.

Vaughn, S. R., Ridley, C. A., & Cox, J. (1983). Evaluating the efficacy of an interpersonal skills training program with children who are mentally retarded. *Education and Training of the Mentally Retarded, 18*(3), 191-196.

Voeltz, L. M. (1984). Program and curriculum innovations to prepare children for integration. In Certo, N., Haring, N., & York, R. (Eds.), *Public school integration of severely handicapped students* (pp. 155-183). Baltimore: Paul H. Brookes.